View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Illinois Digital Environment for Access to Learning and Scholarship Repository

ESSAYS ON MARKET STRUCTURE AND PRICE IN THE KOREAN RETAIL GASOLINE MARKET

BY

MOONSIK KIM

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Seung-Hyun Hong, Chair
Professor In-Koo Cho
Professor George Deltas
Assistant Professor Eun Yi Chung

ABSTRACT

This study empirically examines the relationship between market structure and prices in the Korean retail gasoline market. Korea has experienced big changes in the number of gas stations. For example, nationally the number of gas stations increased 24.9% from 10,406 in 2001 to 13,003 in 2010, but in Seoul the number of stations dropped 24.9% from 816 in 2001 to 613 in 2013.

In the literature, the empirical research about the relationship between the market structure and prices in the retail gasoline market has been done in two ways. First, studies with regression models have mainly focused on finding market structure as the determinant of retail prices. Second, a few studies with structural models have assessed the impacts of mergers on prices. However, previous studies had some limitations: (1) most regression analyses did not evaluate the long-term impact because they used cross-sectional data or short-term panel data; more importantly, they did not successfully correct an endogeneity problem because they used a controversial instrumental variable; (2) because quantity data at the station level were rarely accessible, it was not easy to utilize a structural model.

I try to overcome these limitations of the previous studies in two ways. In the first chapter, I run a regression with long-term panel data and a new instrumental variable which has never been used in the literature. In the second chapter, I estimate a structural model without sales data from gas stations to evaluate the impact of changes to the market structure on prices and welfare.

In Chapter 1, I employ monthly data between January 2003 and December 2011 from seven big cities in Korea to estimate the effects of station density on average retail prices and average sales per station. Instead of population density, which has been commonly used in previous studies, I use the density of diesel cars as the new instrumental variable. The density of diesel cars is a superior instrument to population density in two respects. First, the correlation between station density and the new instrument is obviously expected to be higher than the correlation between station density and population density, because demand for stations depends on the number of cars rather than the number of people. Second, population density may well represent demand for gasoline and therefore have a direct impact on the retail price. Meanwhile, the density of diesel cars satisfies instrument exogeneity because it is not a determinant of gasoline price.

Estimation results yield the following findings. First, regarding a price equation, an OLS estimate of station density is negative but statistically insignificant. However, an IV estimate of station density is negative and statistically significant. The results demonstrate that the OLS estimate underestimates the impact of station density on prices. The IV estimation results show that a 10% increase in the number of stations per square kilometer is associated with a 0.68-0.95% decrease in retail prices. Second, regarding a sales equation, the OLS and IV coefficient estimates of station density are negative, statistically significant and very similar in magnitude. A 10% increase in station density is associated with a 4.2-5.9% decrease in station sales. These findings also imply that the number of stations is an endogenous variable with respect to price, but not with respect to sales. This is the same as the result of Sen and Townley (2010), who examined the retail gasoline industry in Canada.

In Chapter 2, with June 2009 data from 270 gas stations located above the Han River in Seoul, Korea, I use a structural model and conduct counterfactual experiments to estimate the effects of market structure on prices and welfare. My research is different from previous studies. I estimate the model without quantity data from the gas stations by employing the idea of Thomadsen (2005), while previous studies used sales data by directly following the work of Berry et al. (1995). In general, it is difficult to obtain quantity data from gas stations because they keep them secret. Also, data richness allows me to introduce different contractual forms between refiners and stations in the supply model, while previous studies assumed because of data limitations that there was only one type of vertical relationship.

The counterfactual experiments yield the following results. First, although the change to company-owned GS (or HD) stations increases their prices, the prices of company-owned SK stations decrease, decreasing average price. Consumer welfare also decreases because the base utility of GS (HD) stations is smaller than that of SK stations. Second, the change to non-company-owned SK stations decreases the average price because the decreasing effects of ownership changes on prices outweighs the increasing effects of change in vertical contracts on marginal costs, which increases consumer welfare. Third, the change to non-company-owned S-Oil stations greatly lowers the average price mainly because the refiner's wholesale price is the lowest among all refiners. Finally, the exit of stations causes an increase in average price.

The above analysis implies that drop in the number of gas stations leads to very different effects on price and welfare, depending on how brand and contractual form change. Therefore, policies to affect market structure should be developed and implemented with caution because they may have unanticipated effects.

To my parents and family

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Seung-Hyun Hong. His guidance and encouragement helped me throughout the time of researching and writing this thesis.

I would also like to express my appreciation to the rest of my thesis committee: Prof. In-Koo Cho, Prof. George Deltas and Assistant Prof. Eun Yi Chung, for their valuable advice and comments.

This thesis also could not have been completed successfully without data from the Korean Fair Trade Commission and the Korean National Oil Corporation.

Last but not least, I would like to thank my family for all of their love. My father-in-law and mother-in-law always gave me strong support. My father and mother gave me their patient wisdom to overcome many hardships. My wife Hyelim Shin encouraged, supported and loved me. My children, Doyeon Kim and Ryan Kim, have been the pleasure in my long journey.

TABLE OF CONTENTS

CHAPTER 1

GAS STATION DENSITY AND PRICE COMPETITION IN THE KOREAN RETAIL GASOLINE MARKET

1.1 Introduction

In this chapter, I empirically evaluate the effects of the number of gas stations on the average retail gasoline prices and average sales per station. Korea has experienced big changes in numbers of gas stations. For example, nationally the number of gas stations increased 24.9% from 10,406 in 2001 to 13,003 in 2010, but in Seoul the number dropped 24.9% from 816 in 2001 to 613 in 2013. The effects on market performance of changes to market structure are an interesting issue not only to economists but also to policy makers, including competition authorities. For instance, the authorities often block mergers of retail outlets or approve them on condition of the divestiture of some outlets because they believe that mergers may lessen competition and increase prices.

The empirical research about the relationship between market structure and price in the retail gasoline market has been done in two ways. First, studies with regression models have mainly focused on the market structure as the determinant of the retail prices at gas stations, at the city or state level. Second, a few studies with structural models have assessed the impacts of mergers on prices. However, previous studies had some limitations: (1) most regression analyses did not evaluate the long-term impact because they used cross-sectional data or short-term panel data; more importantly, they did not successfully correct the endogeneity problem because they used a controversial instrumental variable; (2) it was difficult to use structural models because the sales data of individual gas stations were not generally available.

I try to overcome these limitations in two ways. In this chapter, I run a regression with long-term panel data and a brand new instrumental variable. In Chapter 2, I estimate a structural model without sales data from the gas stations to examine the impacts of changes to market structure on prices and welfare.

I employ monthly data between January 2003 and December 2011 from seven big Korean cities in to estimate the effects of station density on average retail prices and average sales per station. Instead of population density, which has been commonly used in previous studies, I use the density of diesel cars as the new instrumental variable. The density of diesel cars is a superior instrument to population density in two respects. First, the correlation between station density and the new instrument is obviously expected to be higher than the correlation between station density and population density because demand for stations depends on the number of cars rather than the number of people. Second, population density may well represent demand for gasoline and therefore have a direct impact on retail price. Meanwhile, the density of diesel cars satisfies instrument exogeneity because it is not a determinant of gasoline price.

Estimation yields the following findings. First, regarding a price equation, the ordinary least squares (OLS) estimate of station density is negative but statistically insignificant.

However, the instrumental variable (IV) estimate of station density is negative and statistically significant. A comparison of the two estimates implies that OLS underestimates the impact of station density on price. The IV estimation results associate a 10% increase in the number of stations per square kilometer with a 0.68-0.95% decrease in retail prices. Second, regarding a sales equation, the OLS and IV coefficient estimates of station density are negative, statistically significant and very similar in magnitude. A 10% increase in station density is associated with a 4.2-5.9% decrease in station sales. These findings also imply that the number of stations is an endogenous variable with respect to price, but not with respect to sales, which is the same result Sen and Townley (2010) got when they examined the retail gasoline industry in Canada.

The rest of this chapter is organized as follows: Section 2 discusses the relevant literature; Section 3 describes data and summary statistics as well as the retail gasoline industry in Korea; in Section 4, I provide a theoretical framework; in Section 5, I present regression models and estimation results; Section 6 tests robustness; and Section 7 concludes the chapter.

1.2 Literature Review

This study is partially motivated by the fact that existing theoretical models do not agree in the direction of the relationship between the number of firms and prices. In Section 4, which discusses the theoretical framework, I review in detail the literature about price-increasing and price-decreasing competition.

There are a number of empirical studies which use regression models to address the relationship between the number of sellers and prices in the retail gasoline industry.[1,2] As mentioned in the Introduction, they mainly use cross-sectional data or short-term panel data because they focus on finding what factors are determinants of retail prices rather than evaluating the long-term impact of station density on those prices.

Most studies have found a negative relationship by OLS estimation. Barron et al. (2004) used a one-day price survey at the station level in four U.S. cities (Phoenix, Tucson, San Diego and San Francisco). They found that a 50% increase in the number of stations within a 1.5 mile radius meant a 0.3-0.6% decrease in the average price, depending on the city. Meerbeeck (2003) used weekly prices at the station level in Belgium from March 1998 to March 2001. However, the data did not contain any variation in the number of stations over time, even though it was three years. He also found a negative relationship between the number of stations located in a municipality and the prices. Pennerstorfer (2009) used cross-sectional data at the station level on 400 Lower Austrian stations in September 2003. He usds a spatial lag model which considered price dependence among stations in the market and also found a negative relationship between the number of stations within 15.5km and prices. Shepard (1993) used data from a cross-sectional census of gas stations in a four-county area in eastern Massachusetts in 1987. She also showed that prices decrease with the number of the stations in a one-mile radius. Cooper and Jones (2007) examined station-level data from Lexington, Kentucky, during a four-month period from May to August 2001. Their OLS estimates indicated

---

[1] A literature review of papers using structural models is presented in Chapter 2.

[2] See Eckert (2011) for a general review of the empirical literature on the retail gasoline industry.

that one more competitor on the commuter route was associated with a 0.28 to 0.57-cent decrease in price. Yoon and Lee (2008) used cross-sectional data at the station level for 380 stations in Seoul, Korea, in December 2007. They reported that a 20% increase in the number of stations within a district was associated with a 0.18% decrease in prices. Nam and Oh (2010) used cross-sectional data at the station level for 694 stations in Seoul in October 2008. Their estimates suggested that one more station within a 1km radius led to a decrease of 2.6 won per liter in gasoline prices. Meanwhile, Hoseken et al. (2008) did not find any association between the number of stations within a 1.5 mile radius and prices, using the weekly station-level data of 272 stations in northern Virginia from 1997 to 1999. As Evans et al. (1993) claimed, however, these studies with OLS regression models for price with market structure variables yielded biased estimates because performance feeds back into structure, causing a simultaneous equation bias.

Unlike the above studies, Clemenz and Gugler (2006) used cross-sectional data at the district level in Austria, with population density as an instrumental variable for station density. Their very similar OLS and IV estimates associated a 10% increase in the station density of districts with a 0.35% average decrease in prices, implying that station density is an exogenous variable with respect to price.

To my knowledge, Sen and Townley (2010) is the only paper on the retail gasoline industry that has explicitly evaluated the long-term effect of station density on prices (and station sales) and used an instrumental variable to avoid biased coefficient estimates of station density. They were motivated by Canada’s rationalization, or significant drop in the number of gas stations, in the 1980s and 1990s. Indeed, the stations declined 44.7% from 23,952 in 1980 to 13,250 in 2000. Therefore, Sen and Townley (2010) examined the impacts of outlet rationalization on retail prices and sales, employing monthly city-level data from 10 Canadian cities between 1991 and 1997. Like Clemenz and Gugler (2006), they also used population density as an instrumental variable. They found that a 10% reduction in station density was associated with a 7.0-8.7% increase in retail prices and an 8.5% increase in station sales. However, as they acknowledged, population density could be a poor instrument because it may well represent demand that has a direct impact on retail prices.

1.3 Retail Gasoline Industry in Korea and Data

1.3.1 Institutional Details

Gasoline is provided to final consumers in three ways. The first way is that refiners supply it directly to consumers who regularly need large amounts of it. Such consumers include airlines and the Department of Defense; in 2007, 11.7% of domestically consumed gasoline was provided in this way. The second way is that refiners sell it to stations, and the stations provide it to the consumers; in 2007, 44.4% of domestically consumed gasoline was provided in this way. The last way is that wholesalers buy it from the refiners and sell it to the stations, who provide it to the consumers; in 2007, 43.9% of domestically consumed gasoline was provided in this way.

There are four refiners in Korea, which are SK, GS, HD and S-Oil in order of market share. In December 2012, there were 12,803 stations in the country. The numbers of the stations selling the respective brands were 4296(33.6%), 3164(24.7%), 2345(18.3%) and 1942(15.2%). The independent stations numbered 1056(8.2%). However, the ratio of independent stations was lower in the city than in the country. For example, in Seoul, while the numbers of the stations with the SK, GS, HD and S-Oil brands were 259(41%), 189(30%), 91(14.4%) and 72(11.4%), there were only 20(3.2%) independent ones.

There are four types of vertical relationship between refiners and stations.[3] In the first, the stations are owned by a refining company, and the managers are employed by the company. Therefore, the retail gasoline prices are determined by the refiners. In the second, the stations are owned and operated by a wholesaler who is not a refiner. The wholesaler tends to own more than two stations and buy gas from only one refiner. In the third, open-dealer stations are independently owned and managed with no investment from refiners.[4] The numbers of company-owned, wholesaler-owned and open-dealer stations in Korea were 1117(8.7%), 654(5.1%) and 11,032(86.2%) in December 2012. However, the ratio of company-owned stations is much higher in the city than in the country. For example, the numbers of company-owned, wholesaler-owned and open-dealer stations in Seoul were 198(31.4%), 82(13.0%) and 351(55.6%) in December 2012.

---

[3] For details of contract forms, see Shepard (1993).

[4] The fourth kind of vertical relationship is lessee-dealer stations. Under the contract, the land and facility are owned by the refiner, but the managers are self-employed. Therefore, the retail prices are determined by the managers, but the refiners have the right to control and inspect the quality of the stations. Some rental fees are included in the contracts. The number of lessee-dealer stations is not known, but is considered to be very small.

The gas stations in Korea increased from 10,406 in 2001 to 13,003 in 2010, a 24.9% increase. A well-known factor among the various causes was a change in the law. The Korean government abolished a regulation about a required distance between stations in November 1995.[5] The change in the number of gas stations was different across cities, as discussed in the next section.

1.3.2 Data and Summary Statistics

I used monthly data from seven big cities (Seoul, Incheon, Daejon, Gwangju, Daegu, Ulsan and Busan) between January 2003 and December 2011, resulting in a sample of 756 observations.[6] The location of each city and some statistics such as population and size are shown in Figure 1.1. The monthly wholesale and retail prices of regular grade gasoline were obtained from Korea National Oil Corporation (KNOC). KNOC collects daily retail gasoline prices from all gas stations and computes average monthly prices by cities. It also collects weekly wholesale prices from refiners and wholesalers and converts them into average monthly prices on a national basis. Therefore, unlike retail prices, wholesale prices contain no variation over cities.

The number of stations and sales in each city were obtained from Korea Oil Station Association (KOSA). KOSA reports the number of stations by brands and by vertical relationships

[5] The regulation required stations to have 700m-1km between them in the cities and 2km in the country.

[6] Seven cities are designated by a related law as big cities. When the population of a city surpasses 1 million, the *big city* designation is discussed.

between refiners and stations at the end of each month. From the data, I computed some variables. First, I computed station density by dividing the number of stations in each city by the size of the city. Second, I used the number of gas stations by brand to calculate the Herfindahl-Hirschman index (HHI). Third, I employed the share of company-owned stations among the total number of stations as another variable. Finally, the average sales per station were calculated by dividing the total sales in each city by the number of stations in it.

The unemployment rate, population and city size were obtained from Statistics Korea. The number of diesel cars was obtained from the Ministry of Land, Infrastructure and Transport. Like station density, population density and the density of diesel cars were calculated by dividing the population and number of diesel cars in a city by its size.

Significant cross-city and time-series variation is necessary to identify the impact of the number of gas stations and ensure that effects do not result from unobserved city-specific and time-specific shocks. In this respect, it is suitable to use Korean data. Table 1.1 and Figure 1.2 show how market structure in the retail gasoline industry changed across cities during the sample period (January 2003 to December 2011). For example, the number of stations decreased 14.04% in Seoul, while it increased 38.49% in Gwangju. HHI decreased 9.19% in Gwangju, while it increased 10.7% in Incheon. The share of company-owned stations decreased 61.25% in Gwangju, while it increased 9.93% in Seoul.

Descriptive statistics are presented in Table 1.2. The average city population and size are 3,280,120 inhabitants and 766 km^2, respectively. The average number of stations per square kilometer is 0.57. The mean population per square kilometer is 4804 inhabitants. The average number of diesel cars per square kilometer is 459. The average retail and wholesale prices of a liter of regular-grade gasoline are 1565.5 and 1464.9 won per liter, respectively. Figure 1.3 shows how the average retail and wholesale prices change over time. The correlation between them seems to be very high and indeed the correlation coefficient is 0.99. Finally, the mean of the monthly sales per station is 103,569 liters with a range across cities and over time.

1.4 Theoretical Framework

It is the conventional view that more firms lead to more competitive pricing, resulting in a negative relationship between number of firms and average price. However, the standard view has been challenged by theoretical studies and empirical evidence. In this section, I compare the predicted relationships between number of stations and average price for some models in the literature. My purpose is to provide insight regarding the predictions and set the stage for a subsequent empirical analysis.

1.4.1 Price-Decreasing Competition

Competition among gas stations can be characterized as a monopolistic competition whose two key concepts are product differentiation and free entry. Specifically, a spatial (or location) model is more likely to be suitable for the retail gasoline industry than a representative consumer model, because consumers prefer stations located near them and are willing to pay a premium for them. Probably the most well-known spatial competition model is the circle model of Salop (1979). He shows a kink in the demand curve for a representative firm, suggesting that the demand curve can be divided into the two regions of monopoly and competition. Then he derives a symmetric Nash equilibrium of price p and the number of firms n for the two regions. The notations L and c mean the number of consumers around the circle and the rate at which a deviation from the optimal brand lowers a consumer's utility, respectively. Fixed and marginal costs are denoted by F and m, respectively.

[Equilibria in monopoly region]

$$p_m = m + \frac{c}{2n_m}, \quad n_m = \frac{1}{\sqrt{2}}\sqrt{cL/F} \tag{1-1}$$

[Equilibria in competitive region]

$$p_c = m + \frac{c}{n_c}, \quad n_c = \sqrt{cL/F} \tag{1-2}$$

From equations (1-2), we can derive the negative relationship between the number of firms and average price. For example, an increase in the number of firms induced by either an increase in the market size L or a decrease in the fixed cost F decreases prices.

Intuitively, this relationship makes sense. As the number of firms increases in a given area, the distances between them decrease. It is clear that with spatial competition, prices increase with the distance between stations. Therefore, the average price is expected to decrease as the number of stations increases.

1.4.2 Price-Increasing Competition

Even though the standard precept of economics is that competition will lower prices, some recent empirical studies have suggested otherwise. Ward et al. (2002) found that the entry of private-label products raiseed the prices of name-brand products in the food industry. Using data for the anti-ulcer drug market from 1997 to 1993, Perloff et al. (2005) found that the price of existing brands increaseed when the entrant was relatively different from them. Thomadsen (2007) reported that it was possible for prices to be higher under duopoly than monopoly competition in the fast-food industry.

Several theories have been suggested to explain the positive relationship between price and number of firms. The first is that when all consumers search to find prices, an increase in the number of firms makes it more difficult to find the lowest price in the market, which reduces the consumers' incentive to search. Therefore, this may cause the equilibrium market price to increase with the number of firms (Stiglitz, 1987; Schulz and Stahl, 1996; Janssen and Moraga-Gonzalez, 2004). In the second, Rosenthal (1980) assumes that every firm faces two classes of consumers, a captured loyal group who consume their favorite brands regardless of price and a switching group who consume from the lower-price company. As more firms enter the market, there is less chance of being the low price setter and hence more reason to focus on the captive market with higher prices.[7] Finally, Chen and Riordan (2007, 2008) and Perloff et al. (2005) show that when a new entry entails product differentiation, prices under a duopoly may be higher than those under a monopoly. For example, Chen and Riordan (2008) demonstrated that the price sensitivity effect dominated the market share effect when the consumer values for the two products were drawn from a bivariate exponential distribution.[8]

1.4.3 Price Equation for Empirical Analysis

Besides the number of firms, prices under spatial competition are expected to increase with consumer transport and marginal costs. Moreover, market concentration may affect the retail prices. However, the direction of the impact may be uncertain, as discussed in Clemenz and Gugler (2005). In one case, prices may increase with the degree of concentration for two reasons: (1) in highly concentrated markets, (tacit) collusion is more likely to occur than in markets with fewer competitors; (2) when outlets' nearest competitors are other outlets of the same firm, they are able to charge higher prices. In the opposing case, since market concentration is endogenously determined, firms with multiple outlets may charge lower prices because of low marginal costs.

---

[7] Equilibrium prices exist in mixed strategies.

[8] Duopolists want to reduce their price below the monopoly level in order to recover demand. This is called the *market share effect*. On the other hand, under product differentiation, a duopolist's demand curve may be steeper than a monoplist's. Therefore, consumers are less eager to buy the duopolist's product in response to a price cut. In other words, the duoplist actually wants to raise their price above the monopolist's level. This is called the *price sensitivity effect*. See Chen and Riordan (2008) for details.

To sum up the discussion in this section, the price equation and partial derivatives can be written as follows. Here, S, T, c and H denote density of firms, transport costs of consumers, marginal costs and market concentration, respectively:

$$P = P(S, T, c, H, \ldots) \tag{1-3}$$

$$\partial P / \partial S = ?;\ \partial P / \partial T > 0;\ \partial P / \partial c > 0;\ \partial P / \partial H = ?$$

Equation (1-3) is the basis for the empirical analysis in Section 5.

1.4.4 Sales Equation for Empirical Analysis

A change in the number of gas stations may affect their average profitability. Ideally, I would need diesel sales and revenue from other sources such as a car wash and convenience store as well as the gasoline sales to evaluate profitability. However, I assume that gasoline sales are an important determinant of station profitability, which seems to be a plausible assumption because the number of gasoline cars is bigger than the number of diesel cars.[9] Because total sales of gasoline for each city and each month are accessible, I can construct average sales per station (=total sales of gasoline /the number of stations) and estimate how much they are impacted by station density.

[9] For example, the number of gasoline cars was 4,080,069, while the number of diesel cars was 2,465,366, in the seven cities in 2011.

1.5 Empirical Specifications and Results

1.5.1 Model

The following is a regression model based on equation (1-3) for evaluating the impact of station density on retail prices:

$$\begin{aligned} \ln RP_{it} = {} & \beta_0 + \beta_1 \ln SD_{it} + \beta_2 \ln HHI_{it} + \beta_3 CO_{it} + \beta_4 \ln WP_t + \beta_5 UR_{it} \\ & + \sum CFE_i + \sum YFE_t + \sum MFE_t + u_{it} \end{aligned} \tag{1-4}$$

Similarly, equation (1-5) is used to evaluate the impact of station density on average sales per station:

$$\begin{aligned} \ln Sales_{it} = {} & \beta_0 + \beta_1 \ln SD_{it} + \beta_2 \ln HHI_{it} + \beta_3 CO_{it} + \beta_4 \ln WP_t + \beta_5 UR_{it} \\ & + \sum CFE_i + \sum YFE_t + \sum MFE_t + u_{it} \end{aligned} \tag{1-5}$$

Here, i refers to a city and t to month or year. RP_{it} and $Sales_{it}$ are the average monthly retail price (in won per liter) of regular-grade gasoline and the average monthly sales per station (in kiloliters) in city i at period t. Station density is the number of gas stations per square kilometer.

The effects of retail market concentration on retail gasoline prices need to be controlled for to make sure that the coefficient estimate of station density is unbiased. Market concentration is commonly calculated by the Herfindahl-Hirschman Index (HHI) found in the literature. Ideally, sales data from each station are needed to calculate HHI according to its definition. However, I did not have access to them. Instead, I control for the impact of market concentration on price by using the number of stations by brands and the share of company-owned stations. This approach is reasonable because the concept of market concentration is based on the distribution of firms or ownership. Therefore, HHI_{it} is computed from the number of stations by brands: SK, GS, HD, S-Oil and independents.[10] The term CO_{it} means the share of company-owned stations.

It is also necessary to control for cost shocks and demand shocks to ensure that the coefficient estimate is unbiased. To control for the impact of cost on retail price, I use the average monthly wholesale price of regular-grade gasoline. As explained above, wholesale prices WP_t vary only over time. Of course, the coefficient estimate of WP_t is expected to be positive. The unemployment rate UR_{it} is used to control for demand shifts resulting from business cycles. An increase in unemployment is expected to lower price as demand for gasoline decreases.

Moreover, I use city fixed effects CFE_i, yearly fixed effects YFE_t and monthly fixed effects MFE_t to control for the effects of unobserved city-specific and time-specific factors

[10] For simplicity of computation, independent stations are regarded jointly as one new brand.

that might also have an influence on retail price. Finally, note that I use the natural log of retail price, station density, HHI and wholesale price to make it easy to interpret the coefficient estimates.[11] To check robustness, alternative specifications including log-linear, linear-log and linear models are used later.

1.5.2 Endogeneity and an Instrumental Variable

I am interested in evaluating the impact of station density on retail price. However, there is a possibility of simultaneous bias in the coefficient estimate of station density. For instance, higher retail prices or sales could lead to entry by new firms. Conversely, lower prices or sales could result in the exit of gas stations. If this is the case, an ordinary least squares (OLS) estimate will be biased.

Moreover, equation (1-4) might omit some variables that are correlated with station density and are determinants of retail price. For example, Eckert and West (2004) showed that a maverick firm, ARCO, entered the market in Vancouver, Canada, and charged lower retail price than competitors to secure a market share.

Following Evans et al. (1993), I use a combination of instrumental-variable (IV) and fixed-effect procedures to correct for the endogeneity issue. As discussed above, many previous studies used OLS estimation despite the endogeneity problem, likely because it was difficult to obtain an appropriate instrumental variable. Clemenz and Gugler (2005), and Sen

---

[11] The share of company-owned stations and unemployment rate are not log-transformed because they are percentages.

and Townley (2010) used population density as an instrumental variable for station density.[12] The logic is that stations choose places where there are many consumers (Clemenz and Gugler, 2005), or that increased population may imply new urban development and retail networks, including gas stations (Sen and Townley, 2010).

However, population density may be not a valid instrumental variable for station density in two respects. First, it is possible that population density is a weak instrumental variable for station density because demand for stations is more likely to depend on the number of cars than the number of people. A correlation between population and number of cars may be low because the number of cars is affected by factors such as income, but there is a high correlation between the number of cars and the number of stations.[13]

Second, as Sen and Townley (2010) point out, population density could well represent demand that has a direct impact on retail prices and may be, therefore, correlated with the error term u_{it} in equations (1-4) and (1-5). An interesting and possible example is found in Clemenz and Gugler (2005).[14] They say that "entry decisions in rural areas depend even more on population density than entry decisions in more densely populated areas" (p.235) and show that the coefficient estimate of station density on population density is bigger in rural districts. This suggests that an increase in population in big cities may not result in a proportionate increase in the number of stations, but just in an increase in sales per station, because there is not much available land in big cities. Therefore, in big cities like those used in this study,

[12] Anderson and Johnson (1999) also used population as a proxy for number of stations at the city level.

[13] An increase in teens or the senior population does not lead to an increase in the number of cars either.

[14] They use population density as their instrument because they do not differentiate between urban and rural areas.

population density does not seem to be an appropriate instrument for station density, since it may directly affect the price through increased demand per station.

Hence, I chose the density of diesel cars as a new instrument for station density because it solves the problems mentioned above. [15] First, in terms of instrument relevance, the correlation between station density and the new instrument is expected to be higher than that between station density and population, because demand for stations depends on the number of cars rather than the number of people. Second, the density of diesel cars satisfies instrument exogeneity because it is not a determinant of gasoline price, while population may represent demand and therefore have a direct impact on retail prices.

1.5.3 Estimation Results

1.5.3.1 Impact of Station Density on Price

Table 1.3 provides the estimation results for price equation (1-4). Column 1 shows the OLS estimation results. The OLS coefficient estimate of station density is negative but statistically insignificant. However, we cannot conclude that the number of stations per square kilometer does not have an impact on the average retail market price because of the endogeneity issue.

[15] As found in any econometrics textbook, a valid instrumental variable must satisfy the following two conditions: it should be correlated with an endogenous variable (instrument *relevance*); and it should be uncorrelated with an error term, which means it does not directly influence a dependent variable (instrument *exogeneity*).

Columns 2 and 3 show the IV estimation results when the density of diesel cars is used as the instrumental variable. Specifically, column 2 shows the first-stage regression results with station density $\ln SD_{it}$ as the dependent variable. Column 3 contains the corresponding second-stage estimation results. The coefficient estimate of the density of diesel cars in column 2 is positive and statistically significant at the 1% level, which implies that an increase in the number of diesel cars leads to an increase in the number of stations. An F-test result also shows that we can reject the null hypothesis at the 1% level that the coefficient estimate of the instrument is equal to zero.[16] Hence, because of the weak correlation between station density and the instrumental variable, I do not worry about the possibility that the second-stage IV coefficient estimate will be biased and inconsistent.

The IV estimate of station density in column 3 is negative and statistically significant at the 5% level. Furthermore, the IV estimate (-0.0692) is significantly larger than the OLS estimate (-0.0175) in terms of absolute value. This result implies that the OLS coefficient estimate of station density is biased and underestimates the impact of station density on prices; and furthermore, that the number of stations is inversely correlated with the average retail price. The coefficient estimate implies that a 10% increase in the number of stations, which is the average increase rate in the number of stations during the sample period across the seven cities, is associated with a 0.69% decrease in the retail prices. It is also economically significant, even though it seems to be very small. The margin rate in the retail gasoline market in 2010 is about

[16] The F-statistic is 18.77, which is bigger than 10.

5% of the retail price according to the Ministry of Trade, Industry and Energy in Korea.[17] Therefore, a 0.69% decrease in the retail price is equivalent to a 13.84% decrease in the margin rate. Additionally, the statistical test is performed to add confidence to the estimation results. The Hausman test rejects the null hypothesis at the 10% level that station density is exogenous with respect to retail gasoline prices. Therefore, station density is found to be an endogenous variable with respect to price.

For comparison, estimation results with population density as the instrumental variable are given in Table 1.4. The IV estimate of station density in column 3 is negative but statistically insignificant. The difference between the two estimation results stems from whether both instruments are valid. The first-stage estimates in column 2 demonstrate that population density is a weak instrument for station density because an F-test result shows that we cannot reject the null hypothesis even at the 10% level that the coefficient estimate of the instrument is equal to zero.[18]

Why is population density a weak instrument for station density? This is because the correlation between population and the number of cars is very low in some cities. Table 1.5 shows the OLS estimation results when car density is regressed on population density and other variables. Indeed, the coefficient estimate of population density is positive but statistically insignificant even at the 10% level. In addition, Figure 1.4 shows how population and the number of cars change over time in Busan. The number of cars shows an increasing trend, but population shows a decreasing one. Similar trends have also been observed in Daegu and

[17] "How to increase transparency in gasoline market," press release, April, 16, 2011.

[18] The F-statistic is 0.23, which is smaller than 10.

recently in Seoul. Even though it is outside the scope of this study to find out why population is decreasing in these cities, two main factors may explain the phenomenon. First, according to Statistics Korea, these cities have a lower birth rate than other cities or provinces.[19] Second, some people with low income tend to move out of cities because of high housing prices.

In summary, the empirical estimates in Table 1.3 offer strong evidence that the existence of more stations is significantly correlated with lower retail price, even though there are conflicting theoretical models in the literature.

1.5.3.2 Impact of Station Density on Sales

Table 1.6 shows the estimation results of sales equation (1-5). The OLS estimates are given in column 1. The coefficient estimate of station density is negative and statistically significant at the 1% level. Columns 2 and 3 provide the IV estimation results with the density of diesel cars as an instrumental variable for station density. The coefficient estimate of station density is also negative and statistically significant at the 5% level. Interestingly, the coefficient estimates are very similar between OLS and IV. Indeed, the Hausman test does not reject the null hypothesis that station density is an exogenous variable with respect to average sales per station. In the above subsection, station density was shown to be endogenous with respect to price, but exogenous with respect to sales. These are the same findings as Sen and Townley

[19] In 2011, Seoul, Busan and Daegu showed birth rates of 1.014, 1.078 and 1.146, respectively, while the national average birth rate was 1.244.

(2010). It implies that firms are more likely to make entry decisions based on price or margin than on sales.

The OLS and IV estimates imply that a 10% increase in the number of stations results in about a 4-6% decrease in station sales. Therefore, the impact of station density on sales is much bigger than its impact (about 0.7%) on retail prices. This may be understood to suggest that the retail gasoline market shows monopolistic competition in which product differentiation and non-price competition are key characteristics.

For comparison, I estimate the sales equation using population density as an instrument for station density. The results are given in Table 1.7. The IV estimate of station density in column 3 is positive and statistically insignificant. Therefore, I conclude that population density is not a valid instrumental variable for station density, but the density of diesel cars is.

1.6 Robustness Check

1.6.1 Additional Functional Forms

A log-log specification was used in equation (1-4) and (1-5) to estimate the impact of station density on retail prices and station sales. However, the results do not depend on the specific functional form chosen. I experimented with linear, log-linear and linear-log models in terms of two dependent variables (price and sales) and an independent variable (station density). Table 1.8 provides OLS and IV estimates of the price equation based on the different functional forms. The estimation results are found in columns 1 and 2 for Log-Log, 3 and 4 for Log-Linear, 5 and 6 for Linear-Log, and 7 and 8 for Linear-Linear. As we can see, the result does not vary in that the coefficient estimate of station density is negative and statistically significant, and its magnitude is very similar for all four specifications.[20] The OLS and IV estimates of the sales equation with the four functional forms are given in Table 1.9. As in the case of the price equation, the impact of station density on average station sales is quite similar for all specifications.[21]

1.6.2 Alternative Measure of Market Concentration

HHI was used in equations (1-4) and (1-5) to control for the impact of market concentration on retail prices and average sales per station. Table 1.10 shows the estimation results for the price equation when CR1 and CR2 are used as alternative measures of market concentration. CR1 and CR2 are calculated based on the numbers of stations by brands. The results do not change much, even though the coefficient estimates of station density become slightly bigger in terms of absolute value. Similarly, Table 1.11 shows the estimation results for the sales equation with CR1 and CR2 as alternative measures of market concentration. The results do not change here, either.[22]

---

[20] A 10% increase in station density is associated with a 0.68-0.95% decrease in retail prices.

[21] A 10% increase in station density is associated with a 4.1-6.7% decrease in station sales.

[22] However, the IV estimate of station density is negative but statistically insignificant when CR1 is used.

1.6.3 Additional Instrument

Until now, I have mentioned only station density as an endogenous variable and have thus used an instrumental variable for it. However, HHI may be an additional endogenous variable because an increase in price may lead to the entry of stations and therefore affect HHI in equation (1-4). Therefore, I use an instrumental variable for HHI as well as station density. Variation in HHI is affected by horizontal mergers. There was a horizontal merger between two refiners, SK and Incheon Oil, in August 2006. As a result, Incheon Oil stations became SK stations in three cities (Seoul, Incheon and Busan) out of the seven.[23] Therefore, I use a merger dummy as an instrument for HHI that takes the value of 1 from August 2006 to December 2011 in these three cities.

The IV estimation results for the price equation are given in Table 1.12. They demonstrate that the coefficient estimate of station density is negative, statistically significant at the 5% level, and does not change much in magnitude.

1.7 Conclusion

I evaluate the impact of the number of gas stations on retail prices and station sales. Previous studies using regression models had two limitations: they did not evaluate the

---

[23] Incheon Oil did not have stations across all cities because it was a small refiner.

long-term impact because they use cross-sectional or short-term panel data; and they did not successfully correct for the endogeneity problem because they ignored it or used a controversial instrumental variable.

I use monthly panel data for the nine years from 2003 to 2011 from seven cities in Korea. Instead of the commonly used population density, I use the density of diesel cars as a new instrument. The new instrument is superior to the old variable in two respects. First, the correlation between station density and the new instrument is expected to be higher than the correlation between station density and population density. Second, the density of diesel cars satisfies instrument exogeneity because it is not a determinant of gasoline price, while population density may represent demand and therefore have a direct effect on retail prices.

I observe two findings about the estimation results. First, station density is an endogenous variable with respect to retail prices. The OLS coefficient estimate of station density underestimates its impact on retail prices. The IV estimation results show that a 10% increase in station density is associated with a 0.68-0.95% decrease in retail prices. Second, station density is an exogenous variable with respect to station sales. The OLS and IV estimates demonstrate that a 10% increase in station density is associated with 4.1-6.7% decrease in station sales.

Chapter 1 Tables and Figures

Figure 1.1 Statistics in Seven Cities

Table 1.1 Change in Market Structure

City	Year	#Stations	Change	HHI	Change	Share of Non-Open-Dealer Stations	Change
Seoul	2003	755	-14.040	2890.101	1.511	40.927	9.933
	2011	649		2933.778		44.992	
Incheon	2003	356	8.989	2560.598	10.698	29.494	-23.103
	2011	388		2834.520		22.680	
Busan	2003	431	14.617	2610.290	4.382	24.130	-12.753
	2011	494		2724.680		21.053	
Daejon	2003	267	7.865	2521.146	-4.751	29.963	-36.263
	2011	288		2401.379		19.097	
Daegu	2003	420	3.810	2634.694	-3.279	25.238	-15.484
	2011	436		2548.292		21.330	
Ulsan	2003	247	15.385	2423.741	9.682	18.219	7.852
	2011	285		2658.418		19.649	
Gwangju	2003	239	38.494	2639.135	-9.191	39.749	-61.237
	2011	331		2396.564		15.408	

Figure 1.2 Number of Stations Across Cities

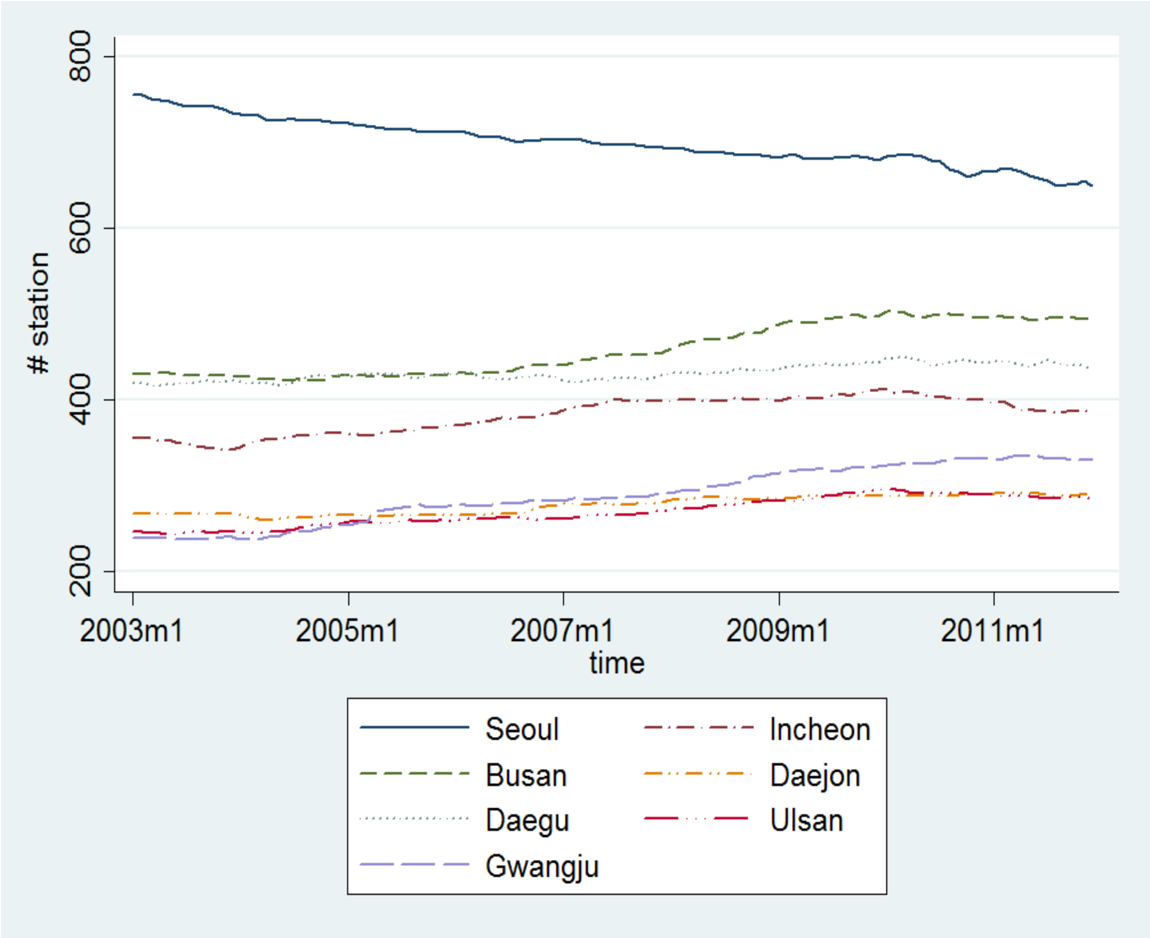

Table 1.2 Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Retail Price (won/liter)	756	1565.497	200.128	1243.630	2053.060
Station Sales (kiloliter)	756	103.569	42.482	55.008	240.108
Wholesale Price (won/liter)	756	1464.912	180.468	1190.770	1890.430
Station Density (#/km^2)	756	0.567	0.267	0.230	1.248
HHI	756	2700.945	204.743	2349.285	3209.829
Share of Non-Open-Dealer Stations (%)	756	29.248	7.786	14.671	49.206
Unemployment Ratio (%)	756	4.022	0.705	1.7	6.5
Density of Diesel Cars (#/km^2)	756	458.803	386.448	97.938	1448.903
Year	756	2007	2.584	2003	2011
Month	756	6.5	3.454	1	12

Figure 1.3 Retail and Wholesale Prices

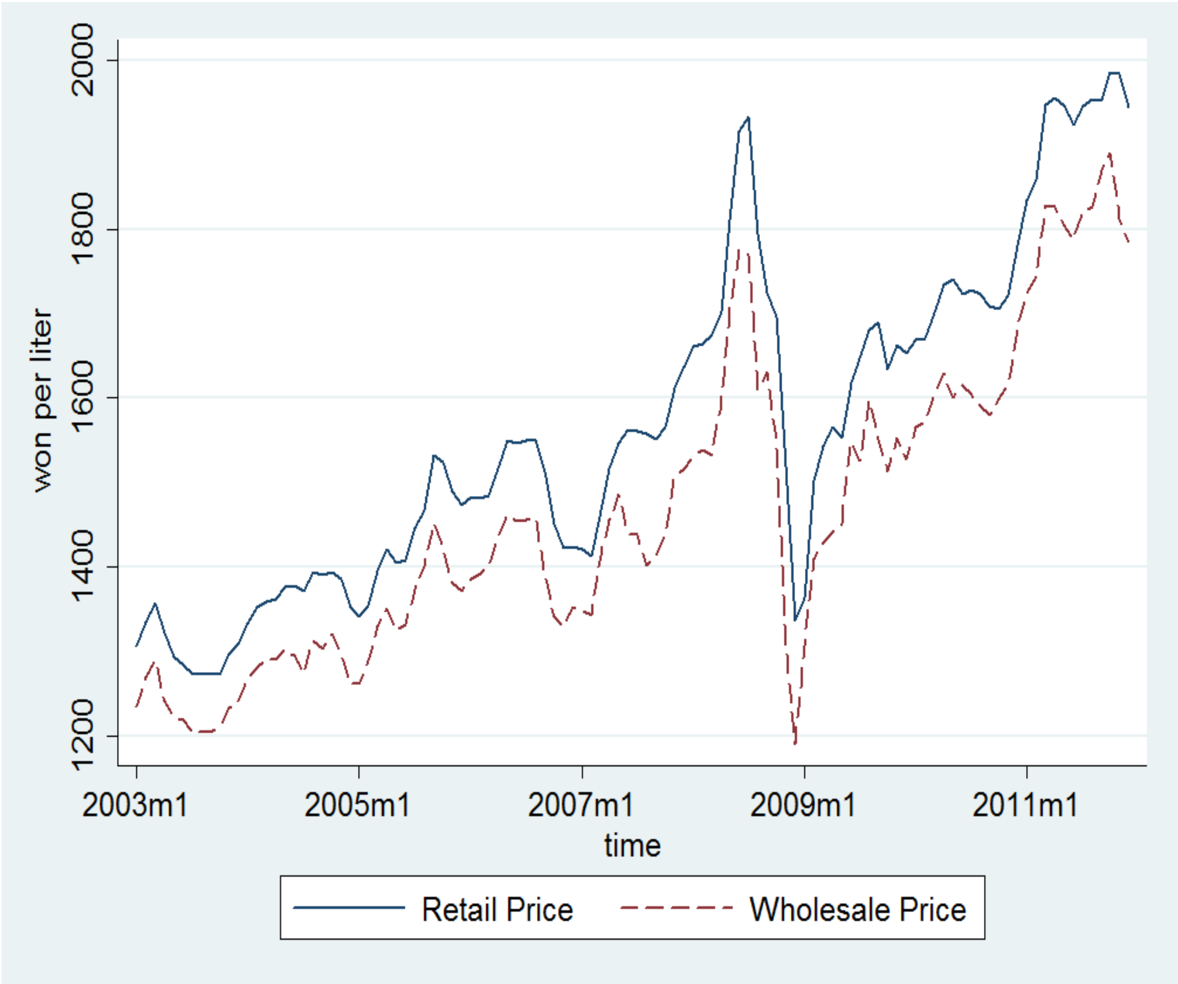

Table 1.3 OLS and IV Estimation Results from Price Equation: Density of Diesel Cars as IV

Variables	(1) OLS Ln(RP)	(2) First Stage Ln(SD)	(3) Second Stage Ln(RP)
Ln(SD)	-0.0175		-0.0692**
	(0.0120)		(0.0337)
Ln(HHI)	-0.0199	-0.0116	-0.0336
	(0.0360)	(0.256)	(0.0278)
UR	-0.00111	-0.00358	-0.00141
	(0.00117)	(0.00440)	(0.00115)
CO	0.000517	-0.00582***	0.000192
	(0.000337)	(0.00184)	(0.000340)
Ln(WP)	0.869***	-0.00250	0.873***
	(0.0113)	(0.0174)	(0.00845)
Ln(Density of Diesel Cars)		1.147***	
		(0.240)	
Constant	1.147**	-7.614***	1.251***
	(0.314)	(1.728)	(0.223)
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Weak IV test	-	18.77 (p=0.004)	-
Hausman Test	-	-	Endogenous (p=0.091)
Observations	756	721	721
R-Squared	0.989	0.995	0.989

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.4 OLS and IV Estimation Results from Price Equation: Population Density as IV

Variables	(1) OLS Ln(RP)	(2) First Stage Ln(SD)	(3) Second Stage Ln(RP)
Ln(SD)	-0.0175		-0.275
	(0.0120)		(0.494)
Ln(HHI)	-0.0199	-0.179	-0.0777
	(0.0360)	(0.318)	(0.103)
UR	-0.00111	-0.00959	-0.00338
	(0.00117)	(0.00607)	(0.00511)
CO	0.000517	-0.00589*	-0.000995
	(0.000337)	(0.00318)	(0.00323)
Ln(WP)	0.869***	0.00347	0.870***
	(0.0113)	(0.0197)	(0.00908)
Ln(Population Density)		0.504	
		(0.961)	
Constant	1.147**	-3.101	1.703*
	(0.314)	(9.200)	(0.941)
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Weak IV Test	-	0.226 (p=0.651)	-
Hausman Test	-	-	-
Observations	756	756	756
R-Squared	0.989	0.992	0.982

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.5 Regression of Car Density on Population Density

Variables	Ln(Car Density)
Ln(Population Density)	0.535
	(0.463)
Ln(Per Capita GRDP)	0.0828
	(0.0710)
Constant	2.922
	(4.532)
City FE	Yes
Year FE	Yes
Month FE	Yes
Observations	756
R-Squared	0.999

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1.4 Population and Number of Cars in Busan

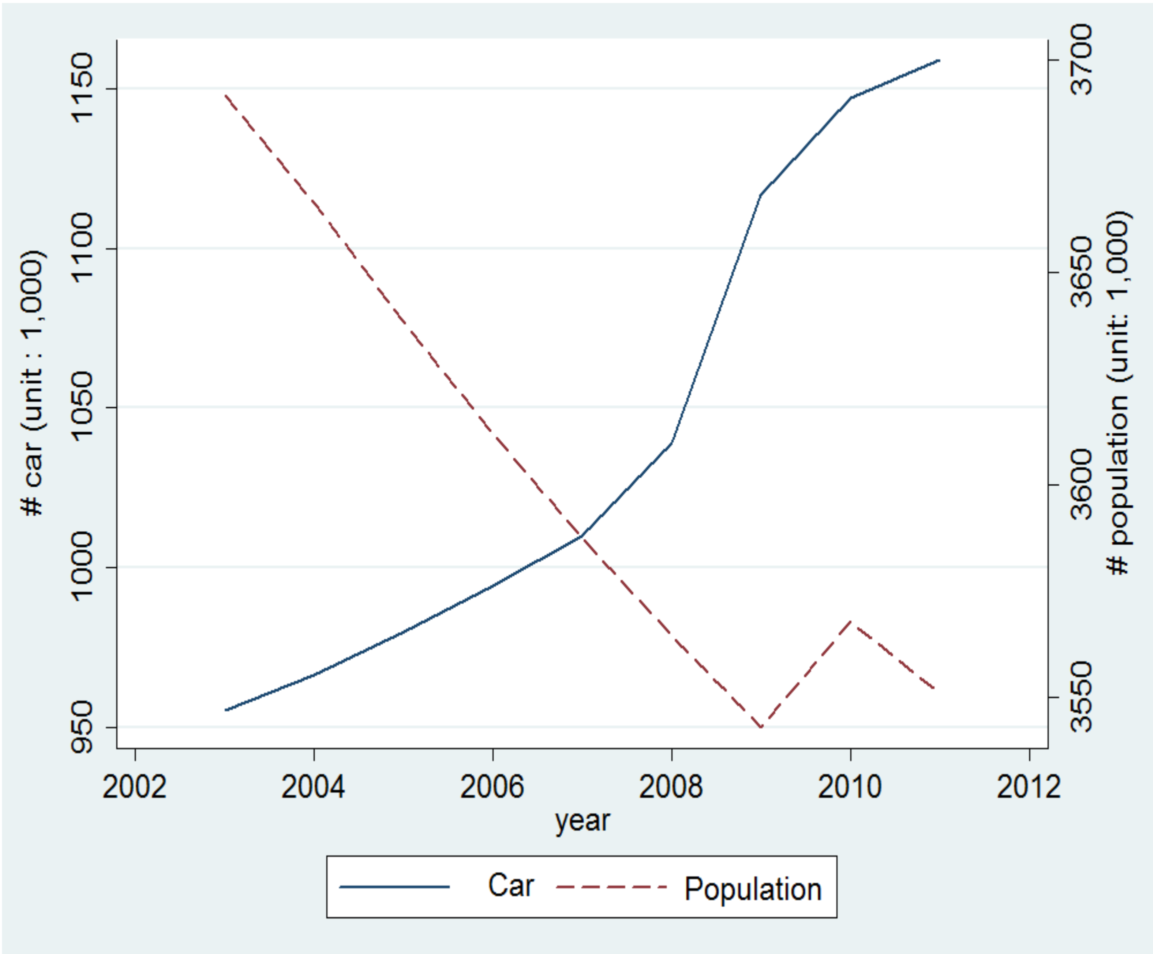

Table 1.6 OLS and IV Estimation Results from Sales Equation: Density of Diesel Cars as IV

Variables	(1) OLS Ln(Sales)	(2) First Stage Ln(SD)	(3) Second Stage Ln(Sales)
Ln(SD)	-0.587***		-0.416**
	(0.104)		(0.178)
Ln(HHI)	-0.452	-0.0116	-0.398**
	(0.239)	(0.256)	(0.198)
UR	-0.0160**	-0.00358	-0.0150***
	(0.00453)	(0.00440)	(0.00484)
CO	-0.00467**	-0.00582***	-0.00360*
	(0.00179)	(0.00184)	(0.00207)
Ln(WP)	-0.422***	-0.00250	-0.421***
	(0.0396)	(0.0174)	(0.0402)
Ln(Density of Diesel Cars)		1.147***	
		(0.240)	
Constant	12.27***	-7.614***	11.76***
	(2.105)	(1.728)	(1.764)
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Weak IV Test	-	18.77 (p=0.004)	-
Hausman Test	-	-	Exogenous (p=0.357)
Observations	756	721	721
R-Squared	0.978	0.995	0.978

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.7 OLS and IV Estimation Results from Sales Equation: Population Density as IV

Variables	(1) OLS Ln(Sales)	(2) First Stage Ln(SD)	(3) Second Stage Ln(Sales)
Ln(SD)	-0.587***		1.242
	(0.104)		(3.834)
Ln(HHI)	-0.452	-0.179	-0.0417
	(0.239)	(0.318)	(0.852)
UR	-0.0160**	-0.00959	7.20e-05
	(0.00453)	(0.00607)	(0.0393)
CO	-0.00467**	-0.00589*	0.00607
	(0.00179)	(0.00318)	(0.0251)
Ln(WP)	-0.422***	0.00347	-0.428***
	(0.0396)	(0.0197)	(0.0537)
Ln(Population Density)		0.504	
		(0.961)	
Constant	12.27***	-3.101	8.324
	(2.105)	(9.200)	(8.106)
F-Test		0.226	
		(p=0.651)	
Hausman Test			-
Observations	756	756	756
R-Squared	0.978	0.992	0.931

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.8 Additional Functional Forms for Price Equation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log-Log		Log-Linear		Linear-Log		Linear-Linear	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
VARIABLES	Ln(RP)	Ln(RP)	Ln(RP)	Ln(RP)	RP	RP	RP	RP
Ln(SD)	-0.0175	-0.0692**			-73.69***	-149.3***		
	(0.0120)	(0.0337)			(16.14)	(29.98)		
Ln(HHI)	-0.0199	-0.0336	-0.0183	-0.0403	-24.61	-46.47	-26.54	-60.88
	(0.0360)	(0.0278)	(0.0380)	(0.0314)	(48.41)	(41.12)	(52.23)	(45.75)
UR	-0.00111	-0.00141	-0.00105	-0.00175	-1.207	-1.647	-1.332	-2.375
	(0.00117)	(0.00115)	(0.00114)	(0.00131)	(1.851)	(1.800)	(1.812)	(1.908)
CO	0.000517	0.000192	0.000562	4.17e-05	0.550	0.0752	0.521	-0.250
	(0.000337)	(0.000340)	(0.000356)	(0.000318)	(0.404)	(0.463)	(0.406)	(0.409)
Ln(WP)	0.869***	0.873***	0.869***	0.873***	1,370***	1,379***	1,370***	1,379***
	(0.0113)	(0.00845)	(0.0114)	(0.00843)	(17.43)	(14.11)	(17.58)	(14.39)
SD			-0.0126	-0.120*			-99.75**	-259.0***
			(0.0210)	(0.0701)			(30.73)	(69.18)
Constant	1.147**	1.251***	1.145**	1.443***	-8,235***	-8,093***	-8,109***	-7,678***
	(0.314)	(0.223)	(0.334)	(0.296)	(407.4)	(316.2)	(460.0)	(384.8)
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Test		18.77		7.82		18.77		7.82
		(p=0.004)		(p=0.031)		(p=0.004)		(p=0.031)
Hausman Test		Endogenous		Endogenous		Endogenous		Endogenous
		(p=0.091)		(p=0.069)		(p=0.034)		(p=0.026)
Observations	756	721	756	721	756	721	756	721
R-Squared	0.989	0.989	0.989	0.989	0.990	0.990	0.990	0.990

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 1.9 Additional Functional Forms for Sales Equation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log-Log		Log-Linear		Linear-Log		Linear-Linear	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Variables	Ln(SALES)	Ln(SALES)	Ln(SALES)	Ln(SALES)	SALES	SALES	SALES	SALES
Ln(SD)	-0.587***	-0.416**			-69.40***	-59.79***		
	(0.104)	(0.178)			(7.909)	(11.00)		
Ln(HHI)	-0.452	-0.398**	-0.475*	-0.438**	-37.75	-34.65**	-41.81*	-40.42**
	(0.239)	(0.198)	(0.239)	(0.203)	(20.81)	(16.96)	(19.84)	(17.23)
UR	-0.0160**	-0.0150***	-0.0173**	-0.0170***	-1.576**	-1.518***	-1.787**	-1.809***
	(0.00453)	(0.00484)	(0.00553)	(0.00642)	(0.488)	(0.471)	(0.563)	(0.576)
CO	-0.00467**	-0.00360*	-0.00509*	-0.00451*	-0.375**	-0.305**	-0.459**	-0.435**
	(0.00179)	(0.00207)	(0.00220)	(0.00272)	(0.115)	(0.134)	(0.172)	(0.205)
Ln(WP)	-0.422***	-0.421***	-0.423***	-0.421***	-42.92***	-42.82***	-43.04***	-42.79***
	(0.0396)	(0.0402)	(0.0395)	(0.0402)	(5.373)	(4.934)	(5.132)	(4.671)
SD			-0.833***	-0.722*			-106.1***	-103.8***
			(0.201)	(0.395)			(11.23)	(23.28)
Constant	12.27***	11.76***	13.38***	12.92***	839.5***	809.7***	992.2***	975.6***
	(2.105)	(1.764)	(2.264)	(2.036)	(168.0)	(136.9)	(162.7)	(143.9)
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Test		18.77		7.82		18.77		7.82
		(p=0.004)		(p=0.031)		(p=0.004)		(p=0.031)
Hausman Test		Exogenous		Exogenous		Exogenous		Exogenous
		(p=0.357)		(p=0.783)		(p=0.418)		(p=0.936)
Observations	756	721	756	721	756	721	756	721
R-Squared	0.978	0.978	0.977	0.977	0.984	0.984	0.984	0.984

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 1.10 Alternative Measures of Market Concentration for Price Equation

	(1)	(2)	(3)	(4)	(5)	(6)
	HHI		CR1		CR2	
	OLS	IV	OLS	IV	OLS	IV
Variables	Ln(RP)	Ln(RP)	Ln(RP)	Ln(RP)	Ln(RP)	Ln(RP)
Ln(SD)	-0.0175	-0.0692**	-0.0245	-0.0862***	-0.0228	-0.0949***
	(0.0120)	(0.0337)	(0.0128)	(0.0334)	(0.0148)	(0.0327)
Ln(HHI)	-0.0199	-0.0336				
	(0.0360)	(0.0278)				
UR	-0.00111	-0.00141	-0.000824	-0.00109	-0.00110	-0.00153
	(0.00117)	(0.00115)	(0.00104)	(0.000988)	(0.00116)	(0.00114)
CO	0.000517	0.000192	0.000496	0.000105	0.000454	-9.37e-06
	(0.000337)	(0.000340)	(0.000283)	(0.000299)	(0.000246)	(0.000257)
Ln(WP)	0.869***	0.873***	0.869***	0.873***	0.869***	0.872***
	(0.0113)	(0.00845)	(0.0110)	(0.00813)	(0.0109)	(0.00805)
CR1			-0.000898	-0.00133***		
			(0.000514)	(0.000296)		
CR2					-0.000538	-0.00111***
					(0.000595)	(0.000419)
Constant	1.147**	1.251***	1.027***	1.040***	1.031***	1.074***
	(0.314)	(0.223)	(0.0780)	(0.0474)	(0.0893)	(0.0532)
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	756	721	756	721	756	721
R-Squared	0.989	0.989	0.989	0.989	0.989	0.989

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.11 Alternative Measures of Market Concentration for Sales Equation

	(1)	(2)	(3)	(4)	(5)	(6)
	HHI		CR1		CR2	
	OLS	IV	OLS	IV	OLS	IV
Variables	Ln(SALES)	Ln(SALES)	Ln(SALES)	Ln(SALES)	Ln(SALES)	Ln(SALES)
Ln(SD)	-0.587***	-0.416**	-0.580***	-0.360	-0.608***	-0.418*
	(0.104)	(0.178)	(0.131)	(0.226)	(0.131)	(0.223)
Ln(HHI)	-0.452	-0.398**				
	(0.239)	(0.198)				
UR	-0.0160**	-0.0150***	-0.0152**	-0.0142***	-0.0164**	-0.0154***
	(0.00453)	(0.00484)	(0.00488)	(0.00518)	(0.00465)	(0.00485)
CO	-0.00467**	-0.00360*	-0.00568**	-0.00421	-0.00596**	-0.00464*
	(0.00179)	(0.00207)	(0.00211)	(0.00272)	(0.00231)	(0.00276)
Ln(WP)	-0.422***	-0.421***	-0.417***	-0.418***	-0.418***	-0.419***
	(0.0396)	(0.0402)	(0.0388)	(0.0387)	(0.0393)	(0.0392)
CR1			-0.00467	-0.00319		
			(0.00478)	(0.00360)		
CR2					-0.00535	-0.00379
					(0.00434)	(0.00340)
Constant	12.27***	11.76***	8.853***	8.711***	9.077***	8.886***
	(2.105)	(1.764)	(0.413)	(0.394)	(0.521)	(0.450)
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	756	721	756	721	756	721
R-Squared	0.978	0.978	0.977	0.977	0.977	0.977

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.12 Additional Instruments for Price Equation

	(1)	(2)	(3)
	OLS	IV Density of diesel cars	IV Density of diesel cars + Merger dummy
Variables	Ln(RP)	Ln(RP)	Ln(RP)
Ln(SD)	-0.0175	-0.0692**	-0.0879**
	(0.0120)	(0.0337)	(0.0358)
Ln(HHI)	-0.0199	-0.0336	-0.0950**
	(0.0360)	(0.0278)	(0.0422)
UR	-0.00111	-0.00141	-0.00142
	(0.00117)	(0.00115)	(0.00111)
CO	0.000517	0.000192	0.000269
	(0.000337)	(0.000340)	(0.000452)
Ln(WP)	0.869***	0.873***	0.872***
	(0.0113)	(0.00845)	(0.00887)
Constant	1.147**	1.251***	1.743***
	(0.314)	(0.223)	(0.351)
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Observations	756	721	721
R-Squared	0.989	0.989	0.989

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter 2

IMPACT OF MARKET STRUCTURE ON PRICES AND WELFARE: EVIDENCE FROM THE KOREAN RETAIL GASOLINE MARKET

2.1 Introduction

In this chapter, I estimate the effects of market structure on prices and welfare using a structural model. The number of retail gas stations decreased by 8.7% from 691 in January 2008 to 631 in December 2012 in Seoul, Korea. The change in the number of gas stations varies by brand and contractual form between refiner and station. The most noticeable feature is that the number of company-owned SK stations decreased by a sharp 36.2% from 116 in January 2008 to 74 in December 2012.[24] Meanwhile, stations with other brands or contractual forms did not decrease and may have increased. It can be inferred that company-owned SK stations exited the market, or changed their brand or vertical contract. Obviously, it can be further concluded that such changes in market structure affected retail gasoline prices.

Motivated by such changes, I evaluate the effects of four possible changes in market structure on gasoline prices and welfare through counterfactual experiments. The changes are (1) company-owned SK stations become company-owned GS or HD stations by changing brand[25]; (2) stations become non-company-owned SK stations by changing contractual form; (3)

[24] SK is the largest refiner in Korea.

[25] There are four brands in the Korean retail gasoline industry: SK, GS, HD and S-Oil.

stations become non-company-owned S-Oil stations by changing both brand and contractual form; and (4) stations exit the market.

There are many empirical studies which evaluate the relationship between market structure and prices in the retail gasoline industry. Most of them use regression models to find the determinants of retail prices such as number of competitors. However, regression results cannot be used to calculate what prices would be under a different market structure because they do not establish a causal relationship between prices and market structure.

A few studies create a structural model and conduct a counterfactual analysis. However, my research is unique in three respects. First, I am interested in various market structure components such as number of stations, brand and contractual form, while previous studies only focused on mergers and computed prices after mergers. Second, I estimate the model without quantity data from gas stations by employing the idea of Thomadsen (2005), while previous studies used sales data by directly following the work of Berry et al. (1995). In general, the sales data of individual stations are rarely accessible because stations keep them secret. Third, data richness allows me to introduce different contractual forms (company-owned and non-company-owned stations) between refiners and stations in the supply model, while previous studies assumed that there was only one type of vertical relationship because of limited data.

The data I use consist of the prices, locations, and station characteristics of 270 gas stations located above the Han River in Seoul, Korea, in June 2009.[26] The identification strategy

[26] Of these 270 stations, 45 are company-owned SK ones.

is based on the work of Thomadsen (2005). A relationship between price and quantity is derived by using the assumption of utility maximization by consumers. The relationship is substituted into the firms' first-order conditions of profit maximization from static Bertrand competition to jointly estimate the parameters of the utility function of consumers and the marginal cost function.

I have some confidence in the estimated model in that (1) the two regression results using original data and fitted data from the structural model are similar, and (2) the own- and cross-price elasticities of stations are realistic in terms of sign and magnitude. Finally, the counterfactual experiments yield the following results. First, the change to company-owned GS (or HD) stations decreases average price. This is because their prices increase, but the prices of the company-owned SK stations decrease. Consumer welfare also decreases because the base utility of GS (HD) stations is smaller than that of SK stations. Second, the change to non-company-owned SK stations decreases average price because the decreasing effects of ownership change on prices outweigh the increasing effects of change in vertical contracts on marginal costs, and increases consumer welfare. Third, the change to non-company-owned S-Oil stations greatly lowers average price mainly because the refiner's wholesale prices are the lowest among refiners, and increases total welfare. Finally, the exit of stations increases average price and decreases total welfare. In particular, I also find that the magnitude of the price increase from the structural model is similar to that from the IV estimation in Chapter 1.

The rest of this chapter is organized as follows. Section 2 discusses the related empirical literature. Section 3 describes the data and the retail gasoline industry in Seoul. In Section 4, I present a regression result as a preliminary analysis. Section 5 explains the structural model and estimation method. In Section 6, I present estimation results and a validity check of them. In Section 7, I conduct counterfactual experiments and discuss the meaning of the results. Section 8 concludes the chapter.

2.2 Literature Review

There have been many empirical studies which examine the relationship between market structure and retail gasoline prices by using regression models at the individual station, city or state level. For studies at an aggregate level, Sen (2003) estimated the impact of local retail market concentration and wholesale prices on average monthly retail prices in 11 Canadian cities between 1991 and 1997. He found that wholesale prices were more important than market concentration in determining retail prices at the city level. Sen (2005) used similar data and concluded that an increase in the market share of independent stations was associated with a decrease in retail prices. Sen and Townley (2010) focused on the impact of outlet rationalization on retail prices. They found that a 27% decline in retail gasoline outlets across 10 Canadian cities between 1991 and 1997 resulted in a 9% increase in retail prices.

For studies at the station level, they set up models in which prices (or margins) were a function of station characteristics, local competition, local demography, and vertical structure between refiners and stations. Regarding local competition variables, Barron et al. (2004), Meerbeeck (2003), Clemenz and Gugler (2006), Pennerstorfer (2009) and Shepard (1993) found that the number of stations within a certain radius of a station was inversely associated with price.[27] Barron et al. (2004), Cooper and Jones (2007), Yoon and Lee (2008) and Nam and Oh (2010) found that price increased as the distance to the closest station increased.[28] Regarding station characteristics, Hastings (2004), Meerbeeck (2003) and Nam and Oh (2010) found that the presence of independent stations caused competitors to lower prices.[29] Regarding contractual forms, Shepard (1993) found evidence that company-owned stations charged lower prices for full service unleaded regular (premium) gasoline than lessee-dealer or open-dealer stations by 6 (10) cents.[30]

While regression analyses make a contribution in that they identify the factors that determine retail prices, regression results cannot be used to calculate what prices would be for stations under different market structures because they do not establish a causal relationship between prices and market structure.

Meanwhile, there are not many papers which have estimated structural models in the retail gasoline market because it is difficult to obtain quantity data from gas stations. They have included the distance between consumers and stations in a utility function to consider geographical differentiation between stations and then estimated the demand and supply

---

[27] However, Hoseken et al. (2008) showed no relationship between station density and prices.

[28] However, Hosken et al. (2008) found no association between station prices and distance to closest station.

[29] However, Pennerstorfer (2009), using data on 400 stations in Lower Austria, argued that the competitive effect of unbranded stations may be muted because an increase in the number of independent stations meant fewer branded stations, reducing price competition among branded stations.

[30] However, Hastings (2004), Hosken et al. (2008) and Nam and Oh (2010) did not find that an increase in the number of company-owned stations affected local prices.

model of Berry et al. (1995),[31] largely in order to calculate the price impacts of mergers.[32] Manuszak (2010) used the monthly data of open-dealer stations on two Hawaiian islands (Maui and Kauai) between 1990 and 1995.[33] He examined the effects of hypothetical mergers of upstream petroleum companies on the retail prices and welfare of refiners, stations and consumers. For example, a Chevron-Texaco merger in July 1995 was predicted to cause a 2.6 cents/gallon increase in the retail prices of the affiliated stations and a 0.3 cents/gallon increase in those of non-affiliated stations. Houde (2012) used the bimonthly data of stations in Quebec, Canada, between 1991 and 2001.[34] He specified commuting paths as the locations of consumers because stations along driving routes of consumers required the same shopping costs. He showed that an Ultramar-Sunoco merger in Quebec in January 1997 was predicted to cause a 0.38 cents/liter increase in the prices of the competitors near the Sunoco stations, which was similar to the results from a difference-in-difference retrospective analysis.

My research is more similar to Thomadsen (2005) in that I estimate demand and supply models without quantity data from stations. Thomadsen (2005) used cross-sectional data from 100 McDonald and Burger King outlets in Santa Clara County, California, in 1999. He found that the impact of mergers on prices decreased as the distance between the merging outlets increased. However, he used cross-sectional data because hamburger prices are

[31] Davis (2006) also used the structural model based on geographical differentiation in the movie theater industry.

[32] See Whinston (2007) and Budzinski and Ruhmer(2009) for survey papers about merger simulation.

[33] The data do not include company-owned stations.

[34] He assumed that refiners could directly control the gasoline prices of stations through nonlinear pricing, which was equivalent to the assumption that all of the stations in Quebec were vertically integrated with refiners.

relatively stable over time, while I use short-term panel data because gasoline prices change frequently at the station level.

2.3 Retail Gasoline Industry and Data

2.3.1 Retail Gasoline Industry

Gasoline is provided to final consumers in three ways. The first way is that refiners directly supply gasoline to certain consumers who usually need a large amount of gasoline. Such consumers include airlines and the Department of Defense. In 2007, 11.7% of domestically consumed gas was provided in this way. The second way is that refiners sell gasoline to stations, and then stations provide gasoline to consumers. In 2007, 44.4% of domestically consumed gas was supplied in this way. The last way is that wholesalers buy gasoline from refiners and sell it to the stations, who sell it to their customers. This accounted for 43.9% of domestically consumed gas in 2007.[35,36]

Until 2002, refiners had set the wholesale gasoline price based on the crude oil price. When the crude price was high, the domestic retail gasoline price might be higher than the

[35] Gas stations may make fake gasoline or purchase from other sources. The Korea Petroleum Quality and Distribution Authority, on average, inspects the quality of gasoline at all stations three times a year. When stations are found to be making or purchasing fake gasoline, they receive fines or business suspensions. In Seoul, the authority found fake gasoline in only 3.1% of inspection cases in 2010.

[36] By a law, branded gas stations must buy the same brand gasoline. But, some stations might buy different brand gasoline when its price is lower. There is no official statistics about the ratio. However, the transaction is illegal and thus the ratio is expected to be low.

international gasoline price. In fact, imported gasoline was 6.3% of total consumption in 2002. At that time they started setting the wholesale gasoline price based on the international gasoline price.[37] As a result, the importing of gasoline almost disappeared. For example, no gasoline was imported between 2006 and 2011.

There are four refiners in Korea: SK, GS, HD and S-Oil, in order of market share. Korea had 12,803 stations in December 2012. The numbers of stations with the SK, GS, HD and S-Oil brands were 4296 (33.6%), 3164 (24.7%), 2345 (18.3%) and 1942 (15.2%), respectively. The number of independent stations was 1056 (8.2%). However, the ratio of independent stations was lower in the city than in the country. For example, in Seoul, while the numbers of the SK, GS, HD and S-Oil stations were 259 (41%), 189 (30%), 91 (14.4%) and 72 (11.4%), that of the independent stations was only 20 (3.2%).

There are four types of vertical relationships between refiners and stations.[38] First, company-owned stations are owned by refiners, and managers are employed by the refiners. Therefore, the retail gasoline prices are determined by the refiners. Second, stations under wholesaler-owned contracts are owned and operated by a wholesaler who is not a refiner. Each wholesaler tends to own more than two stations and buy gasoline from only one refiner. Finally, open-dealer station refiners own and manage their own stations and have no investment from refiners.[39] The numbers of company-owned, wholesaler-owned and

---

[37] The international gasoline price refers to the mean price of oil traded through Singapore as per the data from Platts, a commodity information and trading company.

[38] For details of contractual forms, see Shepard (1993).

[39] There are also lessee-dealer stations. Under the contract, the land and the facility are owned by the refiner, but the managers are self-employed. Therefore, the retail prices are determined by the managers, but the refiners have the right to control and

open-dealer stations in Korea were 1117 (8.7%), 654 (5.1%) and 11,032 (86.2%) in December 2012. However, the ratio of company-owned stations is much higher in the city than in the country. For example, the numbers of company-owned, wholesaler-owned and open-dealer stations in Seoul were 198 (31.4%), 82 (13.0%) and 351 (55.6%) in December 2012.

As mentioned in the Introduction, the number of gas stations in Seoul decreased by 8.7% from 691 in January 2008 to 631 in December 2012 (Figure 2.1). The trend varies with brand and form of vertical contract. Interestingly, most of the reduction results from a decrease in the number of company-owned SK stations. Figure 2.2 shows that the number of company-owned SK stations decreases sharply by 36.2% from 116 in January 2008 to 74 in December 2012, while the number of company-owned GS and S-Oil stations decreases only slightly, and the number of company-owned HD stations increases slightly. Moreover, as shown in Figures 2.3 and 2.4, the wholesaler-owned and open-dealer stations do not show a clear declining trend like the company-owned SK stations, either.

Even though this phenomenon is outside the scope of this study, a decline in margins and characteristics of SK might explain it. In Seoul during the same period, the margins of the gas stations fell because average weekly retail gasoline prices increased by 58.1% from 1349.83 to 2133.96 won/liter, while average weekly wholesale gasoline prices increased by 74.5% from 1144.93 to 1997.45 won/liter.[40,41] SK is also different from other refiners in that it operates

---

inspect the quality of the stations. Some rental fees are included in the contracts. The number of lessee-dealer stations is not known, but is considered to be very small.

[40] During the same period, the weekly prices of Dubai crude oil increased by 261.2% from $38.23 to $138.09 per barrel.

[41] In addition, since January 2012, the Korean government implemented a policy to increase the number of unbranded stations by providing a financial incentive to stations which changed from branded to unbranded stations. The government's purpose was to

various businesses such as real estate and hotels. Therefore, SK has an incentive to change its business from the retail gasoline sector to other sectors when the margins of their gas stations fall below the margins of their other sectors. For example, SK built a 36-story building with residential and commercial units after it shut down a gas station on Yeouido in Seoul in September 2009. It also published plans to build business hotels after shutting down two gas stations in 2013. Meanwhile, open-dealer stations are likely to respond more slowly than company-owned stations, considering initial investment costs and recovery costs carefully.[42]

2.3.2 Data

This study uses short-term panel data from 270 gas stations located above the Han River in Seoul, Korea. Regular gasoline prices, locations, brands, contractual forms and average weekly wholesale prices were collected by Korea National Oil Corporation (KNOC) every Wednesday in June, 2009.[43] Ownership and availability of a car wash service were collected by Korea Oil Station Association (KOSA) in January, 2009, and are assumed to be fixed until June 2009.

Table 2.1 shows summary statistics. The numbers of the SK, GS, HD and S-Oil stations are 127 (47.0%), 78 (28.9%), 32 (11.9%) and 33 (12.2%), respectively.[44] The numbers of the

---

strengthen competition in the retail gasoline market. Industry experts say this policy partly caused the margin of branded stations to decrease and stations to exit the market.

[42] When a station closes down, it is required to clean the site and fill the gas tank with earth.

[43] Wholesale prices are not observed at the station level.

[44] There were no independent stations during the period.

company-owned and non-company-owned stations are 62 (23%) and 208 (77%), and the number of the company-owned SK stations is 45 (16.7%). The average retail price is 1655.1 won/liter and the average wholesale price is 1487.2 won/liter. Specifically, the average retail prices by brand are 1675.0 (SK), 1653.2 (GS), 1633.0 (HD) and 1604.4 (S-Oil). Car washes are found in 178 (65.9%) of the stations. For nine stations, the closest station is owned by the same owner. On average, stations have about 9 competing stations within a 1.5km radius. The average distance to the nearest station is 419.9m. The mean price of the land where a station is located is 4,595,556 won.

I supplement the data of the gas stations with population, workers and car distribution across census blocks. Population is originally collected by census block. The number of workers is collected by administrative districts which are bigger than census blocks, and therefore the number of workers by census block is computed in proportion to its area size within a district. Similarly, the number of cars by census block is calculated in proportion to the number of households within a district. The data of population, workers and cars were obtained from Statistics Korea and Seoul City. The size of the area above the Han River in Seoul is 300.3km,^{2} and there are 8130 census blocks. Table 2.2 provides summary statistics of the demographic data. The average numbers of population, workers and cars across census blocks are 552.9, 234.7 and 128.7, respectively.

2.4 Regression Analysis

Before using the structural model of demand and supply, I set up the following reduced-form model as a preliminary analysis to get a picture of the relationship between market structure and retail prices at a station level:

$$
\begin{aligned}
RP_{it} = \beta_0 + \sum Brand_i + Wash_i + Pump_i + CVS_i + Coown_i + Comp_i + Dist_i \\
+ Popden_i + Workden_i + Carden_i + WP_t + NonCO_i + landP_i + \varepsilon_{it}
\end{aligned} \tag{2-1}
$$

The dependent variable RP_{it} represents retail gasoline prices at station i in week t. The station characteristics variables are brand $Brand_i$, number of pumps $Pump_i$ and dummies indicating whether station i has a car-wash facility or a convenience store, $Wash_i$ and CVS_i. The competition variables are number of competitors within a 1.5 km^2 radius $Comp_i$, distance to the nearest competing station $dist_i$ and a dummy indicating whether the nearest station is co-owned, $Coown_i$. The distance between stations is computed as follows: the addresses of the stations are transformed into their latitudes and longitudes; then the straight-line distance is computed using ArcGIS, a software program for dealing with geographic information in maps.[45]

The cost variables are average wholesale gasoline price WP_t,[46] a dummy for a non-company-owned station $NonCO_i$[47] and the price of the station property $landp_i$. Finally,

[45] For accurate analysis, the actual driving distances should be used. However, I use the straight-line distances as an approximation.

[46] Note that wholesale prices are not observed at the station level.

[47] Company-owned stations avoid double marginalization and thus may charge lower prices than non-company-owned stations.

population density $Popden_i$, worker density $Workden_i$ and car density $Carden_i$ are used to control for local demography. These three density variables are calculated for the nearest census block to station i.

The OLS estimation results are reported in Table 2.3. The signs of most of the coefficients for the variables are the same as would be expected. First, on average, the SK and GS stations charge similar prices, but they are higher than those of the other branded stations. The S-Oil stations are the cheapest. The ranking of the brands by price is similar to that of the brands by number of stations. Stations with car-wash service charge more than those without service. As Yoon and Lee (2008) suggest, stations with a car wash are likely to charge a higher gas price because the service is provided at a low price. Meanwhile, the coefficient estimates of number of pumps and convenience store dummy are not statistically significant even at the 10% level. Therefore, I exclude those two station variables from the structural model later in this chapter.

Regarding the competition variables, prices decrease as the number of competing stations located within a 1.5km radius increase. Moreover, co-ownership increases prices. In other words, stations whose nearest station is also owned by the same owner charge more than stations whose nearest station is owned by a different owner. However, the coefficient estimate of distance to the nearest station is not significant. It may be because I do not include the full layout of station locations in the regression model (2-1).

Regarding the cost variables, wholesale prices and land prices are positively associated with retail prices. However, there is no significant price difference between company-owned stations and non-company-owned stations. This finding is consistent with Hastings (2004) and Nam and Oh (2010), while opposed to Shepard (1993). Prices at company-owned stations may be lower than those at non-company-owned stations because the former buy gasoline at lower wholesale prices than the latter. However, the prices of company-owned stations may be sticky and higher because the refiners, not the individual station managers, make the pricing decisions.

Finally, the estimation results demonstrate that car density for the census block near stations is positively correlated with retail prices. High worker density is also associated with high gasoline prices because the workers may drive their own cars to work and use the surrounding gas stations.

2.5 Model and Estimation

While a regression analysis shows a correlation between market structure and prices, it has a few limitations, as discussed in Thomadsen (2005). First, it is ad hoc to define a relevant geographical market in a regression analysis. For example, Barron et al. (2005) define the market as a circle with a 1.5 mile radius, but it may not be true because the size of the market depends on various factors. Second, it is almost impossible to account for a complicated geographical distribution of stations in a regression. It may take a lot of variables to account for every layout. Finally, we cannot calculate what prices would be for stations under different market structures because a regression model does not specify a causal relationship between prices and market structure.

In order to address these problems, I estimate the structural models of supply and demand. In particular, because the demand model employs station locations, it is not necessary to define the relevant markets, and the data will reveal which stations are competitors.

Basically, I follow the work by Berry et al. (1995) that estimated utilities and costs in differentiated industries from aggregate data because gas stations are geographically and physically differentiated. However, I do not have quantity data because gas stations keep the information secret. Thomadsen (2005) provided insight about how to estimate the model using only price data. A relationship between price and quantity from the utility functions of consumers is derived by the assumption of utility maximization. Then, the relationship is substituted into the firms' first-order conditions of profit maximization. Therefore, the parameters of the utility functions and marginal costs are estimated jointly. However, this approach has an obvious disadvantage. Because quantity data are not used, it is to be expected that I get less efficient estimates than I could obtain if they were.

2.5.1 Demand

A discrete choice framework is used to model demand for gasoline.[48] Consumers choose to either buy gasoline from a particular gas station or use an outside good such as public

[48] Like previous studies such as Manuszak (2010) and Houde (2012), I use the discrete choice model. However, the discrete

transportation. Utility is expected to decrease as consumers travel farther to buy gasoline. Therefore, distance from consumer to station is included in the utility function, which is similar to Thomadsen (2005), Davis (2006), Manuszak (2010) and Houde (2012). Formally, the indirect utility for consumer i from station j is

$$U_{ij} = X'_j\beta - \delta D_{ij} - \gamma P_j + \eta_{ij} \equiv V_{ij} + \eta_{ij}\,, \tag{2-2}$$

where X_j is a vector of dummy variables which indicate brand of station j and whether it has a car-wash[49]; gas price is P_j, and distance between consumer i and station j is D_{ij}. The distance between the consumer's location and the station is also calculated using ArcGIS as the straight-line distance.[50] Therefore, I expect the signs of δ and γ to be positive because of disutility from price and distance. A random utility shock η_{ij} is assumed to follow an i.i.d. type-I extreme value distribution.

When consumer i chooses the outside good $j = 0$, their indirect utility is

choice model does not reflect that the consumer may decrease the amount of purchasing gasoline when the price goes up. Basically, this is the weakness of the model. As an alternative, one may use two-stage approach in order to reflect the continuous consumption within discrete choice framework. First, the consumer chooses a station. Second, he determines the amount of consumption.

[49] For simplicity, I suppress the week subscript t.

[50] For an accurate analysis, the actual driving distance should be used. However, I use the straight-line distance as an approximation.

$$U_{i0} = -\delta D_{i0} + \eta_{i0}, \tag{2-3}$$

where D_{i0} is the distance between the consumer and the outside good, which is considered to be public transportation. The model assumes that consumers who live far from bus stops or subway stations are likely to use their own cars. I assume that D_{i0} is the distance to the nearest gas station, which is similar to Manuszak (2010). This is reasonable because both gas stations and public transportation stops are usually located near crossroads in Seoul.

The previous studies that used the structural model based on Berry et al. (1995) generally included unobserved product characteristics in a utility function, commonly denoted as ξ_j. However, I do not include the term in equation (2-2) for two reasons. First, brand dummies are included and therefore unobserved brand-specific characteristics are controlled for. Second, I assume that consumers' utilities from gas stations depend on observed product characteristics more than unobserved ones such as cleanness of the site or kindness of employees. The assumption seems to be reasonable, considering that 80.6% of respondents selected price and distance to station as the most important factors in choosing stations according to a survey by the Korea Energy Economics Institute (KEEI) in 2010, while only 6.7% selected those unobserved characteristics as the most important factors.[51]

The consumer will choose a gas station which delivers the highest utility if the outside good does not provide a greater utility. Given a utility function and $f(\eta_i)$, the probability

[51] See Kim et al. (2010) for details.

density of the $(J+1)$ dimensional vector η_i, then the share of consumers in location b who buy gasoline from station j is

$$S_{jb}(P, X \mid \beta, \delta, \gamma) = \int_{A_j} f(\eta_i) d\eta_i \,, \tag{2-4}$$

where P is the vector of retail prices at gas stations in the market, and

$$A_j = \{\eta_i \mid (V_{i,j} > V_{i,k} \forall t \neq j) \cap (V_{i,j} > V_{i,0})\} \tag{2-5}$$

is the set of η_i with which a consumer i obtains greater utility in choosing station j than any other station k or the outside good. Because η_i is assumed to follow an i.i.d. type-I extreme value distribution, the share of consumers in location b who choose station j is

$$S_{jb}(P, X \mid \beta, \delta, \gamma) = \frac{e^{X'_j\beta - \delta D_{bj} + \delta D_{b0} - P_j\gamma}}{1 + \sum_{k=1}^{J} e^{X'_k\beta - \delta D_{bk} + \delta D_{b0} - P_k\gamma}} \,. \tag{2-6}$$

Because car owners are consumers to gas stations, I use the number of cars for consumers across census blocks. I also assume that the consumers in each census block are located at the centroid of the census block because census data contains no information about their distribution within the census block. However, this assumption is not too problematic because census blocks are small enough that there are only 128 cars in one census block, on average.

Therefore, the quantity of demand for each station is calculated by summing across all 8130 census blocks, the product of the share of consumers in location b who choose the station and the number of consumers at the location, $h(b)$:

$$Q_j(P, X \mid \beta, \delta, \gamma) = \sum_b h(b) S_{jb}(P, X \mid \beta, \delta, \gamma). \tag{2-7}$$

In addition, the derivative of demand with respect to price is

$$\frac{\partial Q_j(P, X \mid \beta, \delta, \gamma)}{\partial P_k} = \sum_b h(b) \frac{\partial S_{jb}(P, X \mid \beta, \delta, \gamma)}{\partial P_k}. \tag{2-8}$$

2.5.2 Supply

I assume that stations play a static Bertrand game and station owners choose prices at each of their stations in order to maximize the joint profits of all of their stations. This assumption is reasonable, considering that stations can sell as much gasoline as is demanded at the posted prices because gasoline is provided to stations frequently, and prices can be changed at the station level easily and quickly.

Suppose that there are F station owners, each owning a subset F_f of the $j=1,...,J$ stations. The profits to station owner f at week t are

$$\Pi_{ft} = \sum_{j \in F_f} (P_{jt} Q_{jt}(P_t) - c_{jt} Q_{jt}(P_t) - FC_{jt}) . \qquad (2\text{-}9)$$

where c_{jt} is the marginal cost and FC_{jt} is the fixed cost of station j in week t.

I assume that the station's marginal cost is equal to the sum of the cost of the gasoline purchased from the upstream supplier (at wholesale gasoline prices) and a zero-mean, unobserved station-specific portion of the marginal costs. I also assume that wholesale prices at the station level depend on the brand, contractual form and week.[52] Therefore, station j's marginal cost in week t is

$$c_{jt} = \mu_b c_b + \mu_v c_v + \mu_t c_t + \varepsilon_{jt} , \qquad (2\text{-}10)$$

where c_b, c_v and c_t are dummies indicating brands, contractual forms and weeks, respectively. Different branded stations face different wholesale prices because they purchase from different wholesalers. Non-company-owned stations may purchase gasoline at higher wholesale prices than company-owned ones because of double marginalization. Therefore, I introduce contract dummies to verify the possibility in the supply model. For estimation

[52] Transportation cost may affects wholesale prices, but I omit the factor because the impact is expected to be very small.

purposes, vertical contracts are classified into two types (company-owned and non-company-owned) to reduce the number of parameters. The introduction of different contractual forms in the supply model distinguishes it from previous studies. For example, Manuszak (2010) and Houde (2012) used only one type of vertical contract in the supply model because of data limitations.[53] Obviously, wholesale prices also change over weeks, because refiners change prices every week.[54] For estimation purposes, I use a variable for average wholesale price instead of three weekly dummies to reduce the number of parameters. The station-specific portion of marginal costs is attributed to the work efficiency of the employees and the management of the station because gasoline is very similar across all of the stations. Note that $\mu_b c_b + \mu_v c_v + \mu_t c_t$ is independent of ε_{jt} because ε_{jt} depends on the ability and experience of individuals and does not vary over four weeks, while the wholesale price changes every week. As is commonly done in the literature, I assume that station owners know ε_{jt} when they set prices, but the econometrician cannot observe it.

Station owner f maximizes the profit function (2-9) and therefore the first-order condition of station j in week t is

$$Q_{jt}(P_t) + \sum_{r \in F_f} (P_{rt} - \mu_b c_b - \mu_v c_v - \mu_t c_t - \varepsilon_{jt}) \frac{\partial Q_{rt}(P_t)}{\partial P_{jt}} = 0 . \qquad \text{(2-11)}$$

53 Manuszak (2010) used open-dealer stations on two islands in Hawaii. Houde (2012) assumed that refiners could directly control gasoline prices of stations through nonlinear pricing, which is equivalent to the assumption that all stations in Quebec are vertically integrated with the refiners.

54 It is well known that refiners implement weekly pricing in Korea.

For notational convenience, define a matrix Ω as

$$\Omega_{jm} = \begin{cases} \dfrac{\partial Q_m}{\partial P_j} & \text{if } m \text{ and } j \text{ have the same owner} \\ 0 & \text{otherwise.} \end{cases} \tag{2-12}$$

Now, the first-order condition in matrix notation is rewritten as

$$Q(P) + \Omega(P - \mu_b C_b - \mu_v C_v - \mu_t C_t - \varepsilon) = 0\,, \tag{2-13}$$

where $Q(P),\ C_b,\ C_v,\ C_t,\ \text{and } \varepsilon$ are the vector of the quantities of the station, dummies of brand, contractual form and week, and station-specific marginal cost, respectively.

Because I do not have quantity data, I substitute equations (2-7) and (2-8) into equation (2-13), which yields

$$Q(P, X \mid \theta) + \Omega(P, X \mid \theta)(P - \mu_b C_b - \mu_v C_v - \mu_t C_t - \varepsilon) = 0 \tag{2-14}$$

where $\theta' = (\beta', \delta', \gamma')$. Finally, I derive the residual vector for the generalized methods of moments (GMM) estimation:

$$\varepsilon = P - \mu_b C_b - \mu_v C_v - \mu_t C_t + \Omega(P, X \mid \theta)^{-1} Q(P, X \mid \theta)\,. \tag{2-15}$$

2.5.3 Instrumental Variables and Estimation

Given a vector of instrumental variables, Z that are uncorrelated with ε_j but correlated with price P_j, a moment condition for GMM estimation is

$$E[\varepsilon_j(\theta^*) \mid Z_j] = 0 \qquad (2\text{-}16)$$

where θ^* is the true value of θ.

I use brand dummies, a non-company-owned dummy and average wholesale prices as the instruments. They shift the marginal costs of the stations.[55] By assumption, the unobserved component ε_j of marginal cost is not correlated with these three components that make up wholesale prices at the station level. However, the instruments are correlated with the marginal costs of the stations and therefore the prices.

I also use demand shifters as instruments. First, I use the number of competing stations within 1km, 1.5km, …, 3.5km and 4km. The distance to the nearest station is also used as an instrument. Considering the spatial differentiation in the retail gasoline industry, these instruments are consistent with ones suggested by Berry et al. (1995) that are related to characteristics of competing products. Manuszak (2010) and Houde (2012) use similar instrumental variables, even though the distance criteria are different.

[55] They also shift demand.

Second, demographic data are used as instruments. I use the population density and worker density in the nearest census block to the station. These instruments shift demand because sales increase when stations are located near more potential customers.

Third, I use a car-wash indicator as an observed station characteristic. However, I exclude other characteristics such as the number of pumps and a dummy for a convenience store because their coefficient estimates are statistically insignificant in a regression analysis.

Table 2.4 shows the correlation coefficients between retail prices and the instruments. The signs of the correlation coefficients are the same as would be expected. Number of competing stations within a certain radius is negatively correlated with retail prices, while distance to the nearest station is positively correlated. SK's dummy is positively correlated with prices, while S-Oil's dummy is negatively correlated. The correlation between average wholesale prices and retail prices is also positive. However, population density is negatively correlated with retail prices, while worker density is positively correlated. This may not be surprising when we consider the fact that areas with high worker density usually have low population density. The correlation coefficient between the two density measures is indeed negative (-0.1593).

Formally, a set of instruments $Z=(z_1, z_2, \ldots, z_N)$ is used in the estimation. Therefore, the sample analog of the moment condition is as follows.[56]

---

[56] As an alternative, I added a moment condition which requires the sum of the estimated demand for the stations located in each district to match the actual sales data: $G_D(\theta)=\frac{1}{D}\sum_{d=1}^{D}[\sum_{j=1}^{J^d} Q_j^d(\theta)-S^d]$, where d and S denote the district and sales data, respectively. However, the estimation results do not change much and are reported in Table 2.6.

$$G_J(\theta) = \frac{1}{J}\sum_{j=1}^{J} Z_j \varepsilon_j(\theta) \quad (2\text{-}17)$$

The GMM estimator then is the value of $\hat{\theta}$ that solves

$$\arg\min_{\hat{\theta}} \; G_J(\theta)' A G_J(\theta) \quad (2\text{-}18)$$

where A is a weighting matrix for the moments. For the non-linear search for parameters, I use the Nelder-Mead (1995) non-derivative simplex method.[57,58]

2.5.4 Identification

As discussed in Thomadsen (2001) and Thomadsen (2005), the parameters in both the demand and supply models can be separately identified because of geography and the use of both demand and cost shifters as instruments.

Equation (2-14) can be rewritten as

$$P = \mu_b C_b + \mu_v C_v + \mu_t C_t - \Omega(P, X \mid \theta)^{-1} Q(P, X \mid \theta) + \varepsilon .$$

[57] A quasi-Newton method, which is one of the usual search methods, may be used. However, the Nelder-Mead (1995) simplex method is known to be more robust than the quasi-Newton method. See Nevo (2000a) for details.

[58] See Knittel and Metaaxoglou (2008) for examples of searching processes and algorithms.

Therefore, price has two components. The first is the cost component from the supply side that consists of the first three terms on the right side. The component does not vary according to the competitive environment of the station. The second component is the mark-up component from the demand side that consists of the fourth term on the right side. This term varies in different competitive environments because the price elasticity of each station depends on the number of competing stations nearby and how far away they are.

In addition, as mentioned above, the cost shifters and demand shifters are used as the instruments. Therefore, using data from stations that have the same brand and contractual form in a wide variety of competitive environments, I can separate out the constant supply side of the model from the station-specific demand side of the model. In particular, the parameters of the brand dummies on both the demand and supply sides are identified separately.[59]

2.6 Estimation Results and Validity Check

2.6.1 Estimates

[59] For example, if the functional form of the mark-up component is linear in brand dummies, the coefficients of the dummies in both the demand and the cost side could not be separately identified. However, brand dummies in demand side are interacted with price because the mark-up term consists of the multiplication between Ω^{-1} matrix and Q matrix, and each element in Q consists of the multiplication of $e^{X_j'\beta}$ and $e^{-P\gamma}$ as shown in equation (2-6). Therefore, the brand dummies in both the demand side and the cost side can be separately identified.

The estimates of the demand and cost parameters are reported in Table 2.5. First, the p-value of the chi-square test of overidentifying restrictions is 0.78. Therefore, I conclude that the instruments are exogenous and overidentification is not a problem.

Next, the signs and magnitudes of the coefficient estimates seem to be reasonable. For the demand parameters, the coefficients of price and distance from consumers to station are negative and statistically significant at the 1% level, which implies the disutility of price and distance. From the two estimates, the travel cost, which is 646 won/km, is computed by dividing the coefficient of distance by the coefficient of price. The implied travel cost seems reasonable, considering that the bus fare in Seoul was 700 or 800 won in 2009.[60]

For the cost parameters, the average wholesale price coefficient and the estimate of the dummy that indicates a non-company-owned contract are positive and statistically significant at the 1% level. The result makes sense because non-company-owned stations are highly likely to purchase gasoline at higher prices than company-owned stations.

2.6.2 Validity Check of Estimated Model

Before conducting the counterfactual analysis, I check if the estimates are credible in order to have confidence in the estimated model. For this, I do a regression analysis and compute the price elasticities for the gas stations.

---

[60] Some credit cards have the benefit of a 50-won/liter discount for certain branded stations. Considering that consumers, on average, purchase 30 liters at a time, they can get a 1,500-won discount per purchase. They usually drive more than 1km to find a station where they can get such a discount. Therefore, the travel cost seems reasonable.

2.6.2.1 Regression

In section 2.4, a regression analysis was performed with actual prices. Here, I run a regression of the predicted prices from the structural model on the same regressors. If the estimates from the two regressions are similar in terms of magnitude and statistical significance, I can have confidence in the estimated model.

The OLS estimation results are given in Table 2.7. Columns 1 and 2 show the estimation results from the original and predicted prices, respectively. Two results are noted. First, the coefficients estimates from the predicted prices are similar to or at least of the same order of magnitude as those from the original prices. Second, the coefficient estimates of many of the variables are still statistically significant, even though the degree lessens. For example, thirteen estimates are statistically significant when actual prices are used as the dependent variable, while nine estimates are statistically significant when predicted prices are used. The slight decline in the degree of statistical significance results from differences between the original prices and the fitted prices, which may be due to the fact that quantity data are not used in the model.

Hence, I have some confidence that the estimated model is not too problematic.

2.6.2.2 Elasticity

I calculate own-price elasticities and cross-price elasticities for nine gas stations located in a district (Jung-gu) of Seoul. Table 2.8 provides the price elasticities. A map of the stations is shown in Figure 2.5. The nine stations are located along a big street (Toegye-ro) which has six lanes. The own-price elasticities for all of the stations are negative and range from -3.53 to -3.88. The results are consistent with Wang (2009) who found that own-price elasticities for eight gasoline stations ranged from -3.23 to -7.43 in the Perth, Australia, metropolitan area. In addition, the cross-price elasticities between stations are positive and decrease as the distance between them increases, as expected. For example, as found in Figure 2.5, the station in the first row in Table 2.7 is the left-most one on Toegye-ro. The cross-elasticities decrease as the distance from that station to the eight other stations increases, as shown in the table.

The signs and the magnitudes of the elasticities seem to be realistic. Therefore, I have some confidence in the estimated model.

2.7 Counterfactual Analysis

The sharp decline in the number of company-owned SK stations may imply four possible scenarios. First, they may have changed brands, which means that they became company-owned GS or HD stations.[61] Second, they may have changed vertical contracts, which means that they became non-company-owned SK stations. Third, they may have changed both brands and contracts, which means that they became non-company-owned GS or HD stations.

[61] S-Oil does not have company-owned stations.

Finally, they may have exited the market. Motivated by such changes, I use the estimation results to conduct four types of counterfactual experiment and to estimate the impact of the change in market structure on the market prices and welfare.

A vector of new equilibrium prices P^{cf} is obtained under counterfactuals by solving the following equation, given the coefficient estimates from the demand and supply models.[62]

$$P^{cf} = mc^{cf} - \Omega^{post}(P^{cf}, X \mid \hat{\theta})^{-1} Q(P^{cf}, X \mid \hat{\theta}), \tag{2-19}$$

where $mc \; (= \hat{\mu}_b C_b + \hat{\mu}_v C_v + \hat{\mu}_t C_t)$ represents a vector of the marginal costs implied by the estimates from the structural model.

One of the advantages of the structural model is that I can calculate the changes in consumer welfare and producer welfare. Welfare change under the counterfactual is computed using the data from the fourth Wednesday in July, 2009. The change in consumer welfare is measured by a compensating variation. Small and Rosen (1981), Nevo (2000), and Knittel and Metaxoglou (2008) show that the compensating variation for consumer i is given by

$$CV_i = \frac{\ln \sum_{j=0}^{J} \exp(V_{ij}^{post}) - \ln \sum_{j=0}^{J} \exp(V_{ij}^{pre})}{-\gamma}, \tag{2-20}$$

[62] See Knittel and Metaxoglou (2011) for the actual computation process using Matlab.

where V_{ij} comes from equation (2-2).

Therefore, the change in total consumer welfare is computed by

$$\sum_{b} h(b) CV_{ib} \,, \tag{2-21}$$

where CV_{ib} is the compensating variation for consumer i located in census block b. For the welfare change on an annual basis, I assume that consumers purchase 30 liters of gasoline at a time and visit gas stations four times a month, based on Kim et al. (2010).

The change in viable profits for each station owner f is calculated by

$$\prod_f (P^{cf}, mc^{cf}; \hat{\theta}) - \prod_f (P^0, mc^0; \hat{\theta}) \,, \tag{2-22}$$

where P^0 is the initial set of prices. Therefore, the change in total producer profits is given by

$$\sum_f [\prod_f (P^{cf}, mc^{cf}; \hat{\theta}) - \prod_f (P^0, mc^0; \hat{\theta})] \,. \tag{2-23}$$

Finally, the change in total welfare is calculated by summing up the change in total consumer welfare and the change in total producer profits.[63]

---

[63] See Petrin (2002) for an example of quantifying the effects of new products on prices and welfare.

2.7.1 Change in Brands

The first counterfactual is that company-owned SK stations become company-owned GS or HD stations. The change is similar to the assumption that the GS or HD companies acquire company-owned SK stations. Therefore, the elements of matrix Ω in equation (2-19) are changed to reflect the change in ownership. In addition, because the brand is changed for the company-owned SK stations, the marginal cost (mc) and station characteristics (X) are also changed in equation (2-19). Table 2.9 shows the effects of brand change on average price and welfare on an annual basis. When company-owned SK stations become company-owned GS (HD) stations, the average price of the stations decreases by 5.33 (10.95) won/liter and the consumer surplus decreases by $111.22 (146.18) million per year. In addition, the viable profits for the producers decrease by $14.99 (20.03) million per year. Therefore, total welfare decreases by $126.21 million per year for GS-SK mergers and by $166.21 million per year for HD-SK mergers.

Because this counterfactual is similar to mergers between competitors, the prices of the company-owned GS (HD) stations increase. However, there are only 8 (9) of them, which is small. Therefore, the effect of price increases on the average price is limited. Meanwhile, the base utility of GS (HD) stations is smaller than that of SK stations, as found in Table 2.5. Hence, the prices of the company-owned SK stations decrease, of which there are 45. The prices of all stations besides company-owned SK, GS and HD stations increase slightly because strong competitors have disappeared. Overall, average price decreases and consumer welfare also decreases.

This case is an interesting one, especially to a competition agency because both average price and consumer welfare decrease. In general, when an agency reviews a merger, it focuses on whether or not the post-merger price increases, because an increase in price is usually considered a decrease in consumer welfare. However, this case suggests that a merger may lead to a decrease in consumer welfare despite a decrease in price, or conversely an increase in consumer welfare despite an increase in price, because a change in consumer utility resulted from a change in product quality.

2.7.2 Change in Vertical Contracts

To see the impact of vertical contracts on prices and welfare, it is assumed that company-owned SK stations become non-company-owned SK stations. Two opposing effects are expected. First, the marginal costs of stations may increase because non-company-owned stations purchase gasoline at higher wholesale prices than company-owned ones. Second, because ownership transfers to an individual, prices may decrease, which is opposite to the effect of a merger on prices. Therefore, we can anticipate differences in price change among the company-owned SK stations according to their geography. The prices of company-owned SK stations located far from other company-owned SK stations will increase because of an increase in marginal cost, while the prices of company-owned SK stations located close to other company-owned SK stations will decrease, or increase less, because of the two opposing effects.

Figure 2.6 shows the locations of the 45 company-owned SK stations. In fact, the prices of the stations, marked as triangles, located far from others increase, while the prices of the stations, marked as circles, located close to others decrease under the counterfactual.

This suggests that vertical and horizontal mergers have different impacts on price. A change in contractual form is similar to a vertical merger for stations located far from others, while it is similar to a horizontal merger for stations located close to others. Horizontal mergers tend to increase price by reducing competition, while vertical mergers may decrease price by internalizing double marginalization.[64]

The overall effects of change in contractual form are given in Table 2.10. Average price decreases by 5.92 won/liter.[65] It means that the decreasing effect of ownership change on price outweighs the increasing effect of marginal cost change on price. Consumer welfare increases by $12.81 million per year, while viable profits decrease by $32.77 million per year. Hence, total welfare decreases by $19.96 million per year.

2.7.3. Change in Both Brands and Vertical Contracts

[64] The Department of Justice, a competition agency in the U.S., gives more weight to expected efficiencies in determining whether to challenge a vertical merger than in determining whether to challenge a horizontal merger (U.S. Department of Justice Merger Guidelines, June 14, 1984).

[65] The prices of 15 of the 45 company-owned SK stations increase, and the prices of 30 decrease. The prices of other stations also change, and the average price decreases overall.

The case where company-owned SK stations become non-company-owned S-Oil stations is examined.[66] The prices of the stations are expected to decrease because (1) the decreasing effect of ownership change on price is bigger than the increasing effect of marginal cost on price, as shown in Section 2.7.2; and (2) the coefficient estimate of the S-Oil dummy for marginal cost in Table 2.5 is negative, which shows that the wholesale price of S-Oil is the lowest among the four refiners.

The prices of most of the company-owned SK stations decrease under the counterfactual. As a result of price competition, the prices of the other stations also decrease. Therefore, there is a big difference between this counterfactual experiment and other counterfactual experiments where only brand or contractual form changes. Under this counterfactual, the prices of most stations decrease, while under the other counterfactuals, the prices of some stations decrease and some increase. Therefore, under this counterfactual, consumer welfare increases the most among the three counterfactual experiments so far.

Obviously, this result supports the conventional belief that a big increase in consumer welfare arises from price competition among stations.

Table 2.11 shows that average price decreases by 29.3 won/liter, which is 1.72% of the average price on the fourth Wednesday in June, 2009. Consumer welfare increases by $381.36 million per year, and producer profits decrease by $27.45 million per year, which yields an increase of $353.91 million per year in total welfare.

[66] The case where company-owned SK stations become non-company-owned GS (or HD) stations is not reported. It is similar to the case where they become non-company-owned S-Oil stations, even though the magnitude of the effects on average price and welfare are different.

2.7.4 Exit from Market

Finally, the case that all company-owned SK stations exit the retail gasoline market is assumed. Because we found a negative relationship between number of stations and average price in Chapter 1, average price is expected to increase.

Table 2.12 shows the impacts of the counterfactual experiment. Average price increases by 9.43 won/liter. Consumer welfare decreases by $144.37 million per year, and producer profits increase by $10.56 million. Therefore, total welfare falls by $133.81 million.

The exit of 45 company-owned SK stations among the 270 total stations equals a 16.67% reduction in the number of stations. It leads to a 0.55% increase in average price on the fourth Wednesday in June, 2009. This is equivalent to the finding that a 10% decrease in station density is correlated with a 0.33% increase in average price.

A 0.33% increase in average price is lower than the result in Chapter 1 where I found that a 10% decrease in the number of stations was associated with a 0.69% increase in average price. There are two factors to explain the difference between the two numbers. First, market concentration was controlled for in Chapter 1, while exit of stations is included not only in number of stations but also in change in market concentration. Second, the impact of station density on average price is expected to be smaller in Seoul than in other cities, because stations in Seoul face a bigger demand due to the high population or car density. Therefore, the effect of station density on price is similar in both the structural and regression models, using an instrumental variable.

The Korean government has implemented a policy to induce the conversion of full-service branded stations to self-service independent stations by giving some financial incentives since January 2012. The policy was originally introduced in order to strengthen competition and lower gasoline price.

Recently, the number of stations in Korea has started to decrease because their margin is dropping as crude oil prices increase.[67] The counterfactual analysis suggests that the conversion policy may lead to unanticipated effects because it can accelerate the exit of stations and thus increase the price.

2.8 Conclusion

Motivated by the decline in the number of company-owned SK stations in Seoul, I evaluate the impacts of change in market structure on price and welfare through counterfactual experiments. Because a regression analysis cannot do the job, I set up the structural model of demand and supply.

However, sales data from gas stations are not accessible. Therefore, I build the model on the work of Thomadsen (2005) instead of Berry et al. (1995). The relationship between price and quantity is derived from utility maximization. Then the relationship is substituted into firms'

[67] The number of stations in Korea reached to 13,003 in Dec. 2010. Since then, it decreased by 2.4% to 12,687 in Dec. 2013.

first-order condition from static Bertrand competition to jointly estimate the parameters of the utility and marginal cost functions.

After checking the validity of the estimated model, I observe the following results through counterfactual experiments. First, the change in company-owned GS and HD stations decreases average price. This is because their prices increase, but the prices of company-owned SK stations decrease. Consumer welfare also decreases because the base utility of GS and HD stations is smaller than for SK stations. The case suggests that both average price and consumer welfare increase or decrease in the same direction after a merger. Second, the change to non-company-owned SK stations decreases average price because the decreasing effect of ownership change on price outweighs an increase in marginal cost. The counterfactual experiment implies that horizontal and vertical mergers have different effects on the post-merger price, which is consistent with the practices of a U.S. competition agency. Third, the change to non-company-owned S-Oil stations greatly decreases average price mainly because the refiner's wholesale price is the lowest among the four refiners. This supports the conventional belief that a big increase in consumer welfare arises from price competition among stations. Finally, the exit of stations increases average price because it reduces competition. In particular, the result from the structural model is similar to the result from the regression analysis using an instrumental variable in Chapter 1.

The counterfactual analysis shows that drop in the number of gas stations leads to very different effects on price and welfare, depending on how brand and contractual form change.

Therefore, it implies that policies to affect market structure should be developed and implemented with caution because they may have unanticipated effects.

Chapter 2 Tables and Figures

Figure 2.1 Number of Gas Stations in Seoul

Figure 2.2 Number of Company-Owned Gas Stations in Seoul

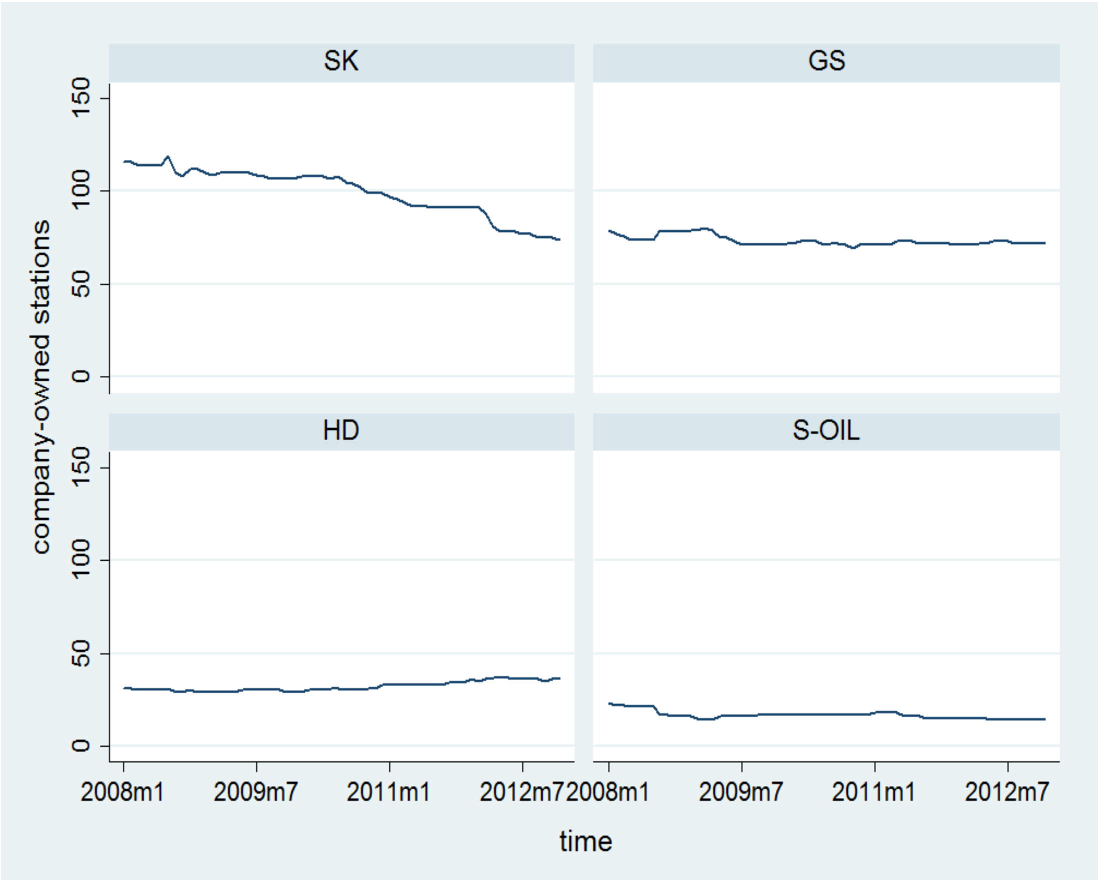

Figure 2.3 Number of Wholesaler-Owned Gas Stations in Seoul

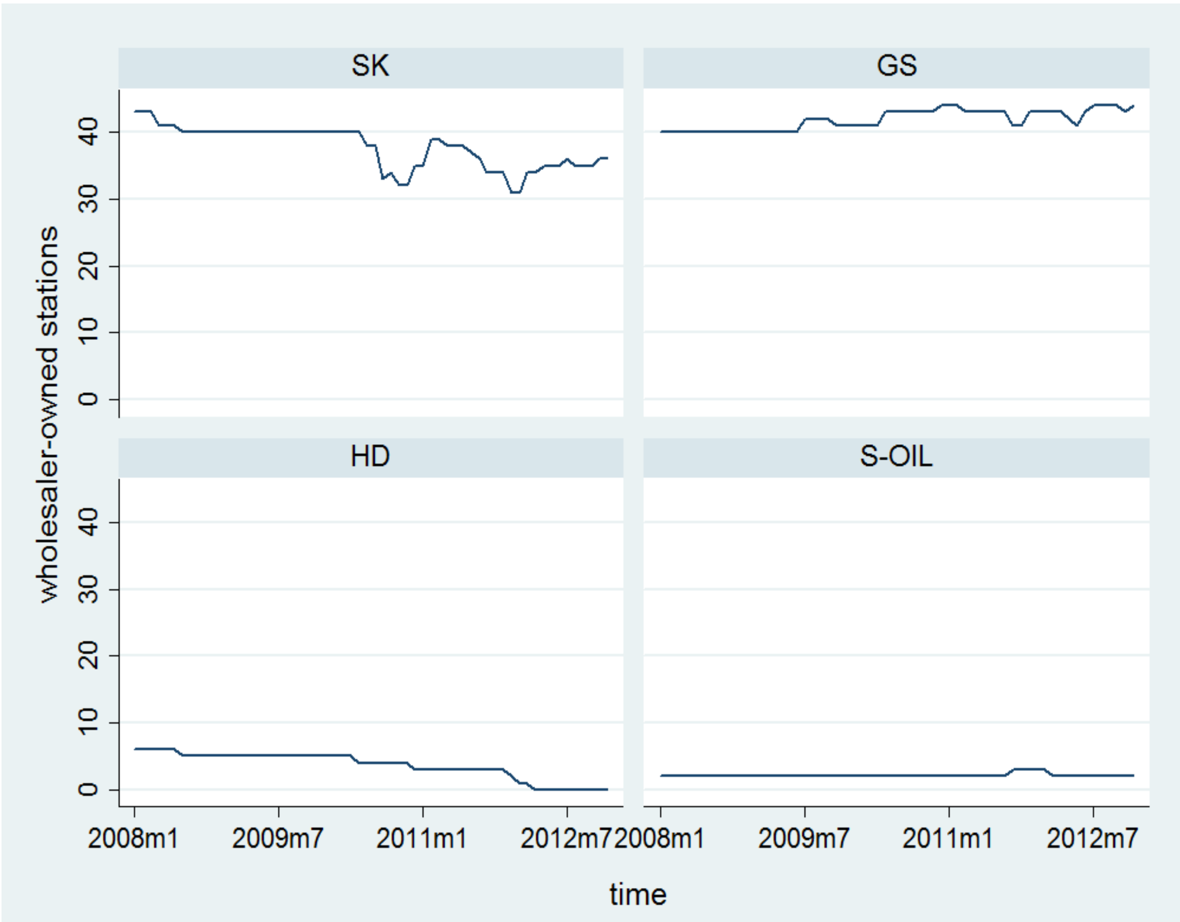

Figure 2.4 Number of Open-Dealer Gas Stations in Seoul

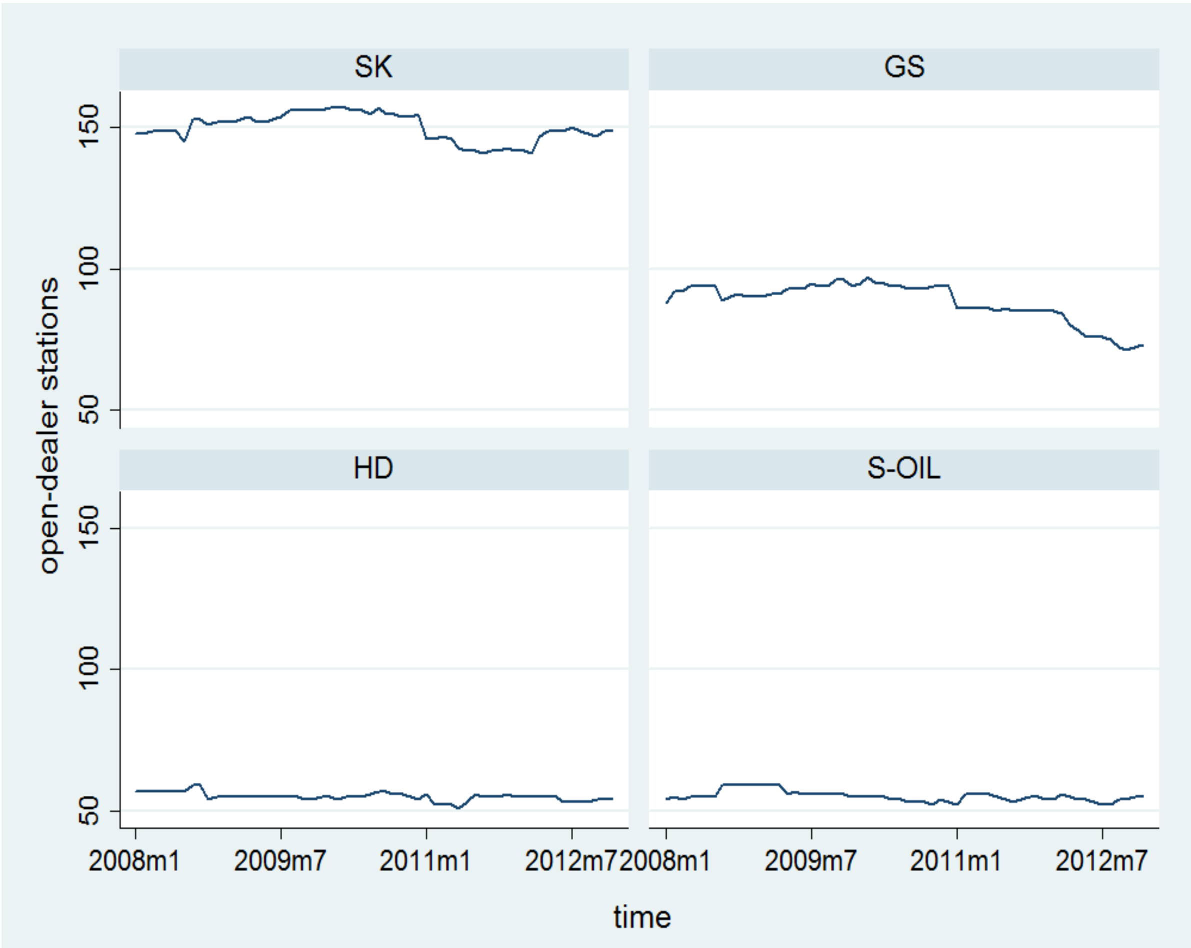

Table 2.1 Summary Statistics of Gas Stations Data

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Retail price	1080	1655.119	88.130	1479	1931
Wholesale Price	1080	1487.233	36.577	1450.65	1543.03
SK Dummy	1080	0.470	0.499	0	1
GS Dummy	1080	0.289	0.453	0	1
HD Dummy	1080	0.119	0.323	0	1
S-Oil Dummy	1080	0.122	0.328	0	1
Car Wash	1080	0.659	0.474	0	1
Non-Company-Owned	1080	0.770	0.421	0	1
Co-Owned Nearest Stations	1080	0.033	0.180	0	1
Number of Stations within 1.5km	1080	9.141	4.183	0	21
Population Density	1080	26.648	23.683	0	151.683
Worker Density	1080	8.302	7.522	0.3028	63.216
Car Density	1080	6.178	4.826	0	29.149
Distance to Nearest Station	1080	419.888	279.332	20.669	2528.594
Land Price	1080	4595.556	2285.484	1110	15300

Table 2.2 Summary Statistics of Demographic Data

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Population	8130	552.878	217.153	0	2543
Workers	8130	234.723	1043.912	0.825	43,801.3
Cars	8130	128.719	63.743	0	751.531

Table 2.3 OLS Estimation Results

Variables	Price
Constant	252.9***
	(41.79)
GS Dummy	-13.72
	(9.616)
HD Dummy	-31.46***
	(11.54)
S-Oil Dummy	-46.47***
	(9.606)
Car Wash	20.68***
	(7.447)
Number of Pumps	0.105
	(0.763)
Convenience Store	-1.032
	(10.80)
Non-Company-Owned	9.960
	(10.82)
Co-Owned Nearest Stations	67.14***
	(14.80)
Number of Stations within 1.5km	-6.840***
	(1.027)
Distance to Nearest Station	-0.00329
	(0.0150)
Population Density	-1.124***
	(0.412)
Worker Density	1.770***
	(0.517)
Car Density	4.487**
	(2.083)
Land Price	0.00809***
	(0.00176)
Wholesale Price	0.945***
	(0.0195)
Observations	1080
R-Squared	0.509

Clustered standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.4 Correlation Coefficients between Retail Prices and Instruments

Instruments	Correlation Coefficients
Number of Competitors within 1km	-0.2945
Number of Competitors within 1.5km	-0.3735
Number of Competitors within 2km	-0.3554
Number of Competitors within 2.5km	-0.3758
Number of Competitors within 3km	-0.3344
Number of Competitors within 3.5km	-0.3232
Number of Competitors within 4km	-0.2992
Distance to the Nearest Station	0.0783
SK Dummy	0.213
HD Dummy	-0.0921
S-Oil Dummy	-0.215
Non-Company-Owned	-0.0348
Wholesale Prices	0.3921
Population Density	-0.1567
Worker Density	0.2521
Car Wash	0.084

Table 2.5 Estimation Results from Structural Model

	Variables	Coefficient Estimates
Demand Side	Constant	8.757 (7.793)
	GS Base Utility	-1.525 (1.176)
	HD Base Utility	-6.973 (5.650)
	S-Oil Base Utility	1.403*** (0.446)
	Price Sensitivity (γ)	-2.308** (1.111)
	Distance Disutility (δ)	-1.492*** (0.254)
	Car Wash	0.213 (0.237)
Supply Side	Constant	0.057 (0.291)
	Marginal Cost GS	0.033 (0.023)
	Marginal Cost HD	0.036 (0.041)
	Marginal Cost S-Oil	-0.137*** (0.039)
	Non-Company-Owned	0.073*** (0.025)
	Average Wholesale Price	0.709*** (0.068)
	Implied Travel Costs (δ / γ)	0.646
	Objective Function (Degree of Freedom)	1.087(3)
	χ^2 p-Value	0.780

(1) SK dummy is used as a base dummy in the demand and supply model.
(2) Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.6 Estimation Results from Structural Model with Additional Moment Condition

	Variables	Coefficient Estimates
Demand Side	Constant	7.821 (6.005)
	GS Base Utility	-1.401 (1.581)
	HD Base Utility	-4.433 (3.656)
	S-Oil Base Utility	1.042*** (0.362)
	Price Sensitivity (γ)	-2.266*** (0.613)
	Distance Disutility (δ)	-1.588*** (0.237)
	Car Wash	0.257 (0.187)
Supply Side	Constant	-0.002 (0.189)
	Marginal Cost GS	0.072*** (0.017)
	Marginal Cost HD	0.057* (0.034)
	Marginal Cost S-Oil	-0.113*** (0.030)
	Non-Company-Owned	0.093*** (0.026)
	Average Wholesale Price	0.711*** (0.082)
	Implied Travel Costs (δ / γ)	0.701
	Objective Function (Degree of Freedom)	3.012(4)
	χ^2 p-Value	0.556

(1) SK dummy is used as a base dummy in the demand and supply model.
(2) Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 2.7 OLS Estimation Results Using Predicted Prices

Variables	(1) Price	(2) Predicted price
Constant	252.9***	-100.3
	(41.79)	(83.03)
GS Dummy	-13.72	-11.59
	(9.616)	(19.68)
HD Dummy	-31.46***	-31.51
	(11.54)	(24.44)
S-Oil Dummy	-46.47***	-38.21*
	(9.606)	(20.36)
Car Wash	20.68***	28.20*
	(7.447)	(15.31)
Number of Pumps	0.105	-0.341
	(0.763)	(1.596)
Convenience Store	-1.032	0.212
	(10.80)	(22.01)
Non-Company-Owned	9.960	12.50
	(10.82)	(21.37)
Co-Owned Nearest Stations	67.14***	57.47*
	(14.80)	(32.94)
Number of Stations within 1.5km	-6.840***	-9.072***
	(1.027)	(2.116)
Distance to Nearest Station	-1.124***	-2.095**
	(0.412)	(0.816)
Population Density	1.770***	3.802***
	(0.517)	(1.058)
Worker Density	4.487**	8.807**
	(2.083)	(4.089)
Car Density	-0.00329	-0.0202
	(0.0150)	(0.0254)
Land Price	0.00809***	0.0130***
	(0.00176)	(0.00348)
Wholesale Price	0.945***	1.172***
	(0.0195)	(0.0408)
Observations	1,080	1,080
R-Squared	0.509	0.326

Clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Figure 2.5 Map of Nine Stations for Elasticity Computation

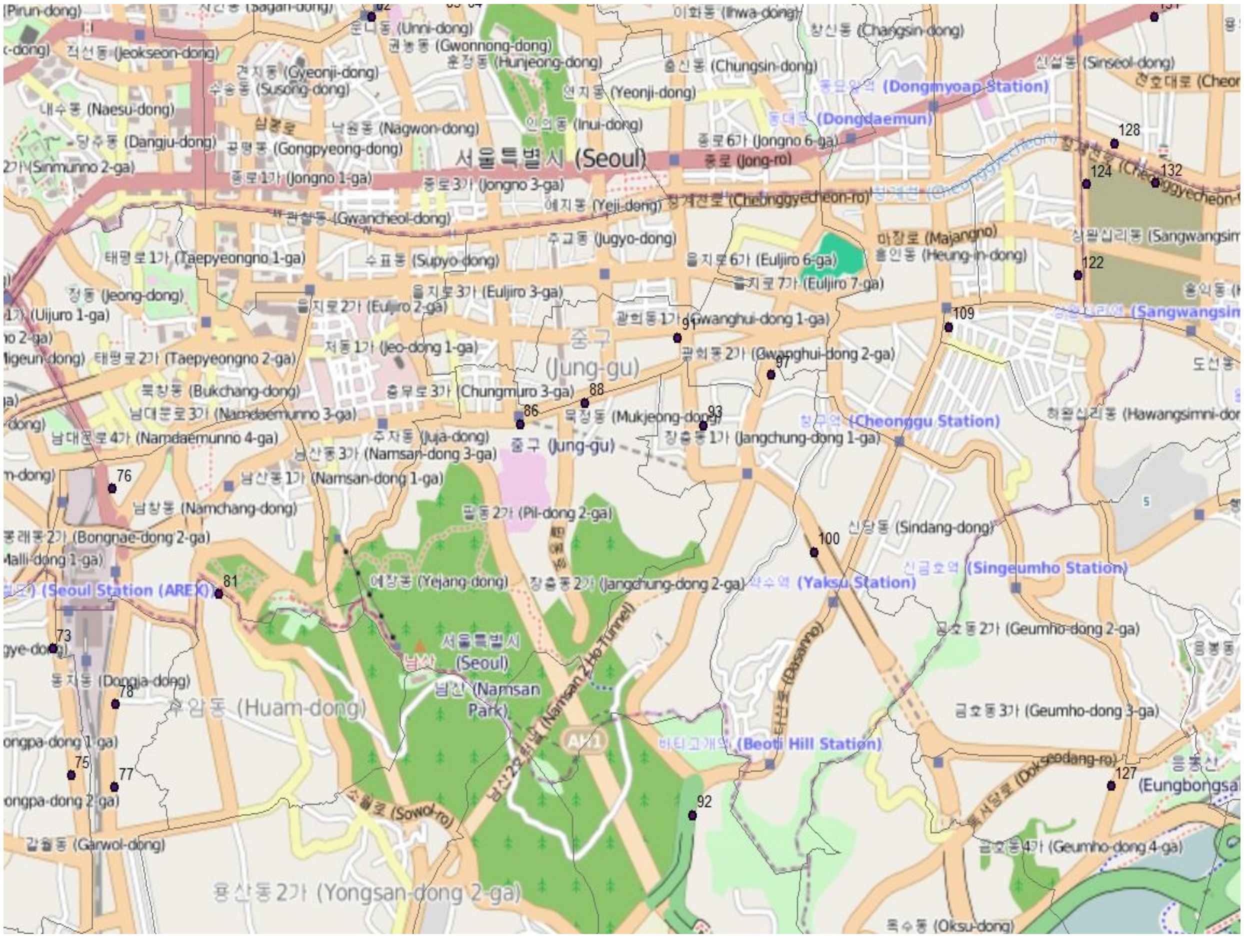

Table 2.8 Own- and Cross-Price Elasticities for Nine Stations at Jung-gu, Seoul

	st_76	st_86	st_88	st_91	st_93	st_97	st_100	st_109	st_122
st_76	-3.826658	0.10053	0.066878	0.037509	0.030607	0.020399	0.014883	0.008066	0.00472
st_86	0.020336	-3.8637	0.059104	0.043687	0.041566	0.032686	0.029616	0.018007	0.011612
st_88	0.081612	0.356546	-3.61606	0.289516	0.281516	0.231871	0.210651	0.136506	0.089661
st_91	0.048038	0.276588	0.303848	-3.71548	0.289853	0.272022	0.232719	0.186088	0.128265
st_93	0.047607	0.319603	0.358819	0.352021	-3.62011	0.353812	0.36193	0.243964	0.165129
st_97	0.033206	0.263036	0.30931	0.345755	0.370295	-3.64452	0.360058	0.298888	0.209926
st_100	0.000031	0.000305	0.00036	0.000379	0.000485	0.000461	-3.88155	0.000366	0.000257
st_109	0.014997	0.165502	0.207974	0.270142	0.291616	0.341364	0.326568	-3.58606	0.345472
st_122	0.007943	0.096605	0.123653	0.16855	0.178671	0.217031	0.207009	0.312722	-3.53228

Terms i and j represent row and column, respectively. Cell entries (i, j) show percentage change in quantity for station j with a 1% change in price for station i.

Table 2.9 Change in Average Price and Welfare from Change in Brands

	Before Counterfactual	After Counterfactual	Change
Average Prices (won/liter)	1698.16	GS: 1692.83 HD: 1687.21	GS: -5.33 HD: -10.95
Consumer Surplus ($Mil)	4422.81	GS: 4311.59 HD: 4276.63	GS: -111.22 HD: -146.18
Variable Profits ($Mil)	818.76	GS: 803.77 HD: 798.73	GS: -14.99 HD: -20.03
Total Welfare ($Mil)	5241.57	GS: 5115.36 HD: 5075.36	GS: -126.21 HD: -166.21

Table 2.10 Change in Average Price and Welfare from Change in Vertical Contracts

	Before Counterfactual	After Counterfactual	Change
Average Prices (won/liter)	1698.16	1692.24	-5.92
Consumer Surplus ($Mil)	4422.81	4435.62	12.81
Variable Profits ($Mil)	818.76	785.99	-32.77
Total Welfare ($Mil)	5241.57	5221.62	-19.96

Figure 2.6 Geography of Company-Owned SK Stations

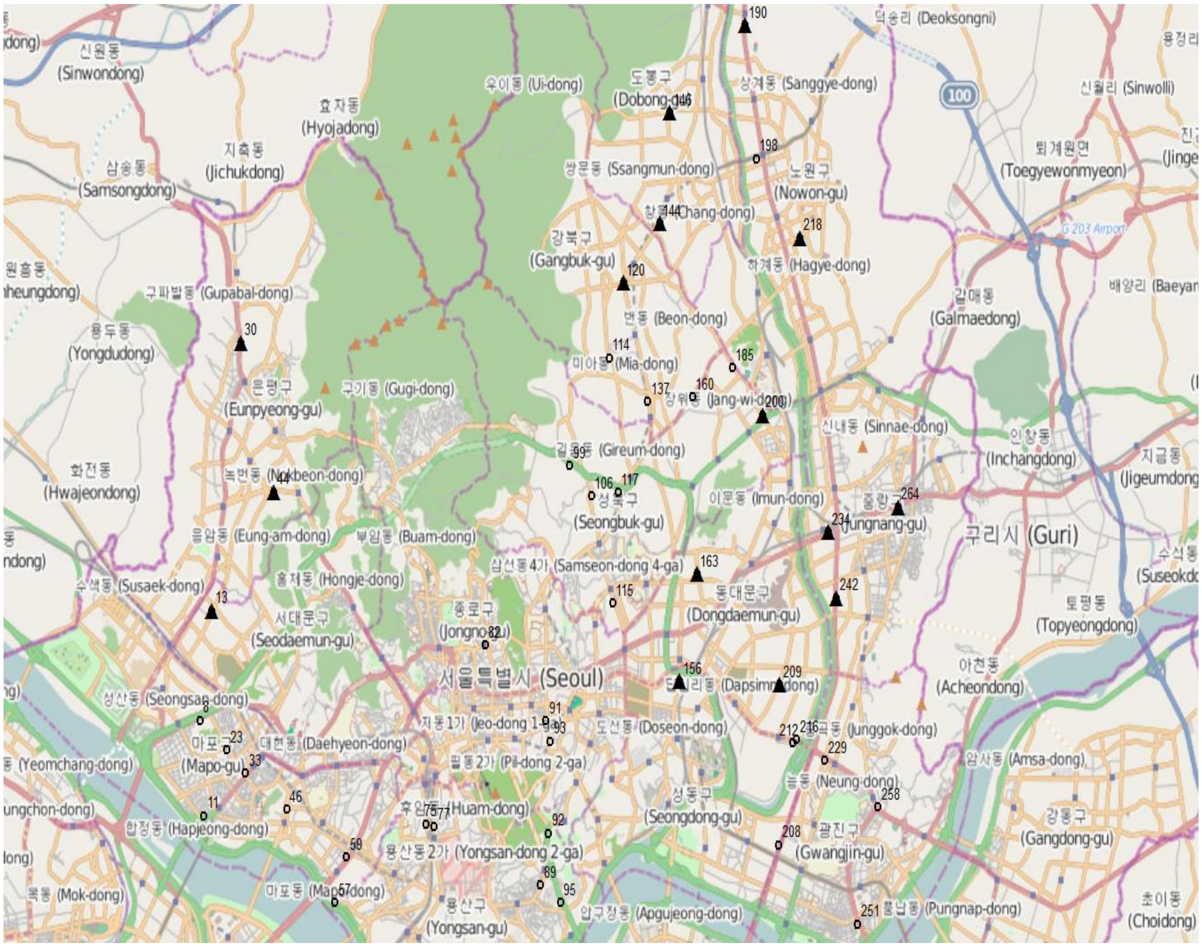

Table 2.11 Change in Average Price and Welfare from Change in Both Brands and Contracts

	Before Counterfactual	After Counterfactual	Change
Average Prices (won/liter)	1698.16	1668.86	-29.3
Consumer Surplus ($Mil)	4422.81	4804.17	381.36
Variable Profits ($Mil)	818.76	791.31	-27.45
Total Welfare ($Mil)	5241.57	5595.48	353.91

Table 2.12 Change in Average Price and Welfare from Change in Number of Stations

	Before Counterfactual	After Counterfactual	Change
Average Prices (won/liter)	1698.16	1707.59	9.43
Consumer Surplus ($Mil)	4422.81	4278.43	-144.37
Variable Profits ($Mil)	818.76	829.32	10.56
Total Welfare ($Mil)	5241.57	5107.75	-133.81

REFERENCES

Anderson, R. W. and R. N. Johnson (1999). “Antitrust and Sales-Below-Cost Laws: The Case of Retail Gasoline.” *Review of Industrial Organization*, 14, 189-204.

Barron, J. M., Taylor, B. A. and J. R. Umbeck (2004). “Number of sellers, average prices, and price dispersion.” *International Journal of Industrial Organization*, 22, 1041-1066.

Berry, S., Levinsohn, J. and A. Pakes (1995). “Automobile Prices in Market Equilibrium.” *Econometrica*, 63(4). 841-890.

Budzinski, O. and I. Ruhmer (2009). “MERGER SIMULATION IN COMPETITION POLICY: A SURVEY.” *Journal of Competition Law & Economics*, 6(2), 277-319.

Chen, Y. and M. H. Riordan (2007). “Price and Variety in the Spokes Model.” *Economics Journal*, 117, 897-921.

Chen, Y. and M. H. Riordan (2008). “Price-Increasing Competition.” *The Rand Journal of Economics*, 39(4), 1042-1058.

Chung, J. W. and J. Y. Lee (2012). *The Study on Price Elasticity of Gasoline and Diesel Demand in Korea*. Seoul: Korea Energy Economics Institute.

Clemenz, G. and K. Gugler (2005). “Locational choice and price competition: some empirical results for the Austrian retail gasoline market.” *Empirical Economics*, 31, 291-312.

Cooper, T. E. and J. T. Jones (2007). “Asymmetric Competition on Commuter Routes: The Case of Gasoline Pricing.” *Southern Economic Journal*, 74(2), 483-504.

Davis, P. (2006). “Spatial Competition in Retail Markets: Movie Theaters.” *The Rand Journal of*

Economics, 37(4), 964-982.

Eckert, A. (2011). “Empirical Studies of Gasoline Retailing: A Guide to the Literature.” *Journal of Economics Surveys*, 1-27.

Eckert, A. and D. S. West (2004). “A tale of two cities: Price uniformity and price volatility in gasoline retailing.” *The Annals of Regional Science*, 38, 25-46.

Evans, W. N., Froeb, L. M. and G. J. Werden (1993). “Endogeneity in the Concentration – Price Relationship: Causes, Consequences, and Cures.” *The Journal of Industrial Economics*, 41(4), 431-438.

Hastings, J. S. (2004). “Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California.” *The American Economic Review*, 94(1), 317-328.

Houde, J. (2012). “Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline.” *The American Economic Review*, 102(5), 2147-2182.

Hosken, D. S., McMillan, R. S. and C. T. Taylor (2008). “Retail gasoline pricing: What do we know?” *International Journal of Industrial Organization*, 26, 1425-1436.

Janssen, M. C. W. and J. L. Moraga-Gonzalez (2004). “Strategic Pricing, Consumer Search and the Number of Firms.” *Review of Economic Studies*, 71, 1089-1118.

Kim, H. G., Won, D. W., Korean Research, and H. M. Gu (2010). *The Survey on Consumers' Purchasing Behavior at Gas Stations*. Seoul: Korea Energy Economics Institute.

Kim, Y. D. (2007). *Petroleum Demand and the Price Elasticity*. Seoul: Korea Energy Economics Institute.

Knittel C. R. and K. Metaxoglou (2008). “Estimation of Random Coefficient Demand Models: Challenges, Difficulties and Warnings.” *NBER Working Paper* No. 14080.

Knittel C. R. and K. Metaxoglou (2011). “Challenges in Merger Simulation Analysis.” *The American Economic Review*, 101(3), 56-59.

Manuszak, M. D. (2010). “Predicting the impact of upstream mergers on downstream markets with an application to the retail gasoline industry.” *International Journal of Industrial Organization*, 28, 99-111.

Meerbeeck, W. V. (2003). “COMPETITION AND LOCAL MARKET CONDITIONS ON THE BELGIAN RETAIL GASOLINE MARKET.” *DE ECONOMIST*,151(4), 369-388.

Nah, I. K. (2001). *The Effects of Petroleum Price Regulation in Korea.* Seoul: Korea Energy Economics Institute.

Nam, J. H. and S. A. Oh (2010). “The Effects of Vertical Relation on Price Competition in Retail Gasoline Markets in Seoul.” *Korean Journal of Industrial Organization*, 18(2), 19-41.

Nelder, J.A. and R. Mead (1965). “A Simplex Method for Function Minimization.” *The Computer Journal*, 7(4), 308-313.

Nevo, A. (2000a). “A PRACTITIONER’S GUIDE TO ESTIMATION OF RANDOM-COEFFICIENTS LOGIT MODELS OF DEMAND.” *Journal of Economics & Management Strategy*, 9(4), 513-548.

Nevo, A. (2000b). “Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry.” *The Rand Journal of Economics*, 31(3), 395-421.

Pennerstorfer, D. (2009). “Spatial price competition in retail gasoline markets: evidence from Austria.” *The Annals of Regional Science*, 43, 133-158.

Perloff, J. M., Suslow, V. Y. and P. J. Seguin (2005). "Higher Prices from Entry: Pricing of Brand-Name Drugs." *Mimeo*, University of California-Berkeley.

Petrin, A. (2002). "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy*, 110(4), 705-729.

Rosenthal, R. W. (1980). "A Model in which an Increase in the Number of Sellers Leads to a Higher Price." *Econometrica*, 48(6), 1575-1579.

Salop, S. C. (1979). "Monopolistic Competition with Outside Goods." *The Bell Journal of Economics*, 10(1), 141-156.

Schulz, N. and K. Stahl (1996). "Do Consumers Search for the Highest Price? Oligopoly Equilibrium and Monopoly Optimum in Differentiated-Products Markets." *The Rand Journal of Eonomics*, 27(3), 542-562.

Sen, A. (2003). "Higher prices at Canadian gas pumps: international crude oil prices or local market concentration? An empirical investigation." *Energy Economics*, 25, 269-288.

Sen, A. (2005). "Does Increasing the Market Share of Smaller Firms Result in Lower Prices? Empirical Evidence from the Canadian Retail Gasoline Industry." *Review of Industrial Organization*, 26, 371-389.

Sen, A. and P. G. C. Townley (2010). "ESTIMATING THE IMPACTS OF OULTET RATIONALIZATION ON RETAIL PRICES, INDUSTRY CONCENTRATION, AND SALES: EMPIRICAL EVIDENCE FROM CANADIAN GASOLINE MARKETS." *Journal of Economics & Management Strategy*, 19(3), 605-633.

Shepard, A. (1993). "Contractual Form, Retail Price, and Asset Characteristics in Gasoline

Retailing." *The Rand Journal of Economics*, 24(1), 58-77.

Small, K. A. and H. S. Rosen (1981). "Applied Welfare Economics with Discrete Choice Models." *Econometrica*, 49(1), 105-130.

Stiglitz, J. E. (1987). "Competition and the Number of Firms in a Market: Are Duopolies More Competitive than Atoistic Markets?" *Journal of Political Economy*, 95(5), 1041-1061.

Thomadsen, R. (2001). *Empirical Studies of Differentiation in the Fast-Food Industry*. Ph.D Dissertation, Department of Economics, Stanford University.

Thomadsen, R. (2005). "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries." *The Rand Journal of Economics*, 36(4), 908-929.

Thomadsen, R. (2007). "Product Positioning and Competition: The Role of Location in the Fast Food Industry." *Marketing Science*, 26(6), 792-804.

U.S. Department of Justice and Fair Trade Commission (2010). *Horizontal Merger Guidelines*.

Wang, Z. (2009). "Station level gasoline demand in an Australian market with regular price cycles." *The Australian Journal of Agricultural and Resource Economics*, 53, 467-483.

Ward, M. B., Shimshack, J. P., Perloff, J. M. and J. M. Harris (2002). "EFFECTS OF THE PRIVATE-LABEL INVASION IN FOOD INDUSTRIES." *American Journal of Agricultural Economics*, 84(4), 961-973.

Whinston, M. D. (2007). "ANTITRUST POLICY TOWARD HORIZONTAL MERGERS." *Handbook of Industrial Organization*, 3, 2371-2440. Amsterdam: Elsevier.

Yoon, Y. H. and E. Y. Lee (2008). "The Analysis of Gasoline Price Determination in Seoul." *Korean Journal of Industrial Organization*, 16(2), 76-94.